Research Note 82-19

RESEARCH AND EVALUATION IN SUPPORT OF AN EXECUTIVE PROGRAM IN NATIONAL SECURITY

BASIC RESEARCH

U. S. Army

Research Institute for the Behavioral and Social Sciences

March 1982

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Note 82-19	2. GOVT ACCESSION NO. AD A117634	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Research and Evaluation in Support of an Executive Program in National Security		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Douglas M. Johnston		8. CONTRACT OR GRANT NUMBER(s) MDA 903-79-G-0004
9. PERFORMING ORGANIZATION NAME AND ADDRESS John F. Kennedy School of Government Harvard University Cambridge, Massachusetts 02138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q161102B74F
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333		12. REPORT DATE March 1982 13. NUMBER OF PAGES 67
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Program evaluation
Senior executive training
General/flag officer training
Decision making
Evaluation strategies
Case study approach
Short-term training
National security

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Harvard executive training program methodology was applied to the development of a short-term training program for senior U.S. national security officials. The model program was assessed for effectiveness in meeting the program objectives. The principal strategy used in evaluation was participants' perceived competence in dealing with seven major national and international security areas where their personal decisions can critically affect political economic or military interests of the United States. Results showed that each

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

group of participants (29 in 1978, 30 in 1979, 79 in 1980) reported substantial gains from pre-program to post-program evaluations, and further gains at 6 months post-program points. (The 1980 course was increased to 80 spaces because of the wide-spread interest generated by the earlier attendees.) Additionally, War College graduates characteristically experienced marginally higher gains than do non-War College attendees. Also, each year's group was brought to essentially the same six-month post-program level of perceived competence, independent of their respective pre-program levels. The course was continued in 1981 and will continue for the foreseeable future.

U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

JOSEPH ZEIDNER
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Research accomplished under contract
to the Department of the Army

John F. Kennedy School of Government
Harvard University

NOTICES

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report summarizes the findings of a three year research program supported by a grant from the U.S. Army Research Institute to assess the effectiveness of short-term training programs for senior national security officials. The research involved a twofold thrust: (1) that required to develop an appropriate model, and (2) that associated with assessing model effectiveness.

The model that was developed, the Harvard Executive Program in National and International Security, was designed specifically to enhance the effectiveness of high level military and civilian personnel who are in (or moving into) posts where their personal decisions or recommendations can critically affect the political, economic, or military interests of the United States. The basic intent behind the program was to provide increased understanding of selected substantive areas and, more importantly, to provide tools that participants would find useful on a continuing basis.

An evaluation of the first three years of the program has resulted in the following conclusions:

- Short-term training along the lines of the Harvard model can effectively raise the levels of participants' perceived competence in critical areas of national security concern.
- Most participants continue to benefit beyond the termination of the program through an ongoing application of newly acquired skills to their daily occupations.

Average "Competence" Across All Areas
(Scale: 100 to 1,000)

Year	Pre-Program	Post-Program	Six Months Post
1978	619	714 (+14%)	741 (+4%)
1979	589	714 (+22%)	752 (+5%)
1980	509	683 (+34%)	733 (+7%)

- Those entering with lower levels of perceived competence tend to gain proportionally more than their colleagues who enter with higher levels of proficiency.

- There has been an increase in value added by the program as the trend has been toward accepting applicants with lower levels of perceived competence.

- The program has brought each year's group to essentially the same six month post-program level of competence, independent of their respective pre-program levels.

- War College graduates characteristically experience marginally higher gains than do non-War College graduates. The same observation holds true for military vs. civilian participants, with the former realizing the greater gains.

INTRODUCTION

This is the third and summary report on the effectiveness of short-term training programs for senior national security officials. In assessing training effectiveness, research was conducted on two fronts. First there was the research required to develop a program that could serve as a suitable model, i.e. the Harvard Executive Program in National Security. Second was that which was required to assess how effectively the program met its basic objectives.

The research to support the program itself primarily involved the development of appropriate case studies. As the program's principal pedagogical instrument, the case study represents one of the more rewarding forms of research, particularly where application is concerned. Exposure to high-level officials provides an immediate and demanding test of accuracy and effectiveness and, through the feedback that ensues, serves to enrich the quality of the case materials. This has certainly proven to be the case with those studies developed to date (see Attachment A).

While the case-related research is a critical ingredient in the makeup of the program, it is the assessment of program effectiveness that constitutes the principal focus of this report. To properly understand the approach taken, a thorough knowledge of the program background becomes helpful.

BACKGROUND

As the panoply of government problems has grown increasingly complex, the need for improved executive development in the public sector has become more readily apparent. The very manner in which the Federal Service is structured, with its continual turnover of political appointees in most of

the top policy jobs (some of whom are qualified; others not) tends to foster inefficiency in policy design and execution.

A recent survey of top level officials within the national security community suggests a strong need to improve management capability in both the military and civilian spheres. On the military side, there is currently very little available in the way of continuing education at the "corporate executive level." While there is a plethora of training programs and opportunities available for military personnel at the 0-4 through 0-6 levels (Lieutenant Commander through Captain; Major through Colonel), there is very little beyond that. Each of the Services does send its Flag officer selectees to an indoctrination course of two or three week's duration, but these courses, for the most part, have an internal Service focus and tend to be somewhat mechanical in nature.

Although little in the way of executive training is offered for Flag and General officers, the need for something substantial is quite real. In some respects, the term "General" ought to equate to "generalist manager." In many instances, however, the path to Flag or General officer rank is quite specialized; and once there, the individual can serve for as many as ten to fifteen years with virtually no opportunity for intellectual or managerial refurbishment. Adding to the problem is the fact that there are certain areas of critical importance to the top Service manager which are not covered well in any of the multitudinous forums to which he or she have been exposed along the way. Here reference is made specifically to how business is conducted between the Department of Defense (DOD) and other agencies in the Executive Branch; the interaction between the Executive departments and the Congress; how the different perspectives of the Office of the Secretary of Defense (OSD), the Joint Chiefs of Staff (JCS), and the

Services come into play; what is involved/required in getting decisions implemented; and the like. The situation is further exacerbated by the inevitable tendency for the incumbent (or the incumbent's superior) to feel that the job demands at that level are too consuming to permit extrication for any significant period of training.

The needs on the civilian side are even more pressing. In May of 1974, the U.S. Army Management Engineering Training Agency (AMETA) conducted a "Study of Management and Executive Development in Industry, Universities, and the Federal Government" (13). In this study, the authors struck a comparison between executive development efforts in the private and public sectors and found the latter wanting. They concluded that executive development in government is anything but institutionalized; that there is little, if any, accountability by senior managers; and that there is no system for identifying, training, and tracking prospective executives.*

More recently, in March of 1976, a panel of the National Academy of Public Administration completed a study entitled "Strengthening Civilian Executive Development in the Department of Defense" (8). Conducted at the request of the Assistant Secretary of Defense (Manpower and Reserve Affairs), the study generated a number of recommendations that accommodated what the panel found at that time to be the general characteristics of Federal Civil Service (i.e., limited lateral entry, little mobility between agencies or occupational groups, and virtually no "conscious system" for developing executive leadership) while specifically addressing the unique problems of DOD. In this regard, the panel saw the need for consistent top

---

* To be sure, the recent advent of the Senior Executive Service with its attendant emphasis on management has helped to some extent; but many of the observations based on the earlier system continue to hold true.

level support of executive development programs as the most critical problem facing the Department. Sheer size coupled with the fact that most career posts in Defense are occupied either by political appointees or the military (with attendant brief periods of incumbency) makes it very difficult to personify the top level interest that is required.

The National Academy study also pointed out that over the five year period from 1976 through 1981, approximately 50% of the DOD supergrade workforce would become eligible for retirement. Thus, it was suggested that a primary focus of DOD executive training should be to develop the replacements for what is likely to be an inordinately high number of vacancies. Related to this suggestion was the fact that the civilian personnel system, while not a closed system, is heavily dominated by single career personnel. Indeed, approximately 80% of DOD supergrade equivalent jobs are filled from within the Department. The inference is twofold: (1) it is unrealistic to shortchange executive development on the basis that potential candidates have already received such training in private industry prior to entering government, and (2) any executive development program for DOD should be heavily slanted toward meeting the specific requirements associated with managing the national security process.

In view of the above, Harvard University decided in April of 1978 to offer a two week Executive Program in National Security on an annual basis, with the first program to take place in August of the same year. In January 1979, the Navy Personnel Research and Development Center published a Navy Civilian Executive Study (8) that implicitly endorsed the intent behind the Harvard effort (and any others like it):

> Although relatively few executives have had extensive academic training in management, leadership, or administration, they spend most of their time performing tasks in these areas. This highlights a major training need. It includes...general management knowledge and skills (e.g. decision-making, communications)...A need was also identified for the integration of civilian and military training in the shore establishment.

PROGRAM GOALS

Modeled after earlier programs pioneered by the Business School for senior business executives, the Executive Program in National Security (subsequently renamed the Executive Program in National and International Security) is designed for Flag and General rank military officers and for high-level civilian officials who are either in, or moving into, posts where their personal decisions or recommendations can affect the critical political, economic or military interests of the United States. Its principal goals are to improve participants' understanding of:

-- differences in interest, perspective and style associated with varying types of responsibilities in differing organizations;

-- the economic, political, technological, and organizational context within which national security policies and programs are framed; and

-- various dimensions of high-level management, including the uses and misuses of formal analysis in decision-making and policy design.

While the principal themes of the program include management perspective, policy development, and uses of analysis, it is the dimension of perspective that constitutes the central thread of the curriculum. As noted by Gordon Pask of System Research Ltd. during a recent ARI-sponsored conference on decision making in complex systems:

> "There are and there must be multiple perspectives in order to have decisions at all. These could be different individuals or perspectives entertained by one individual. Individuals are not very relevant in this sense because, when they engage in activity, perspectives tend to become distributed. Individual differences, if one recognises that they are not context free; are important...Good decisions must be defined in context. Good decision making involves the ability to abduce and resolve juxtaposed abductions, and the ability to maintain the variety of the system while avoiding chaos." (11)

In getting participants to think critically about the perspectives, stakes, and interests of their counterparts in and around government, it becomes necessary to analyze the setting in which problems arise and the particular incentives which various actors face. Individuals behave differently in different organizational settings according to: (1) their own history and that of their organization, (2) the reward/sanction structure posed by their career system, and (3) which of the many faces of an issue they confront. Thus, a key goal of the program is to increase the participants' awareness and understanding of the respective frames of reference of the Presidency; the Congress; the media; political appointees; careerists, whether they be military or civilian; the private sector; and other countries. These categories were chosen as representative of the work areas germane to national security policy development and program implementation.

COURSE DESCRIPTION

The pedagogical approach used in the program is highly participatory in nature. Because those government and private sector executives who attend collectively share extraordinary experience and wisdom, the curriculum is designed so that participants derive maximum benefit not only from the faculty, but from one another. Case studies, supplemented by seminars

and some lectures, provide intense, vicarious experience in a broad spectrum of managerial situations and increased insight into a wide range of important security-related issues. These typically include such topics as global economic forces and the determinants of national economic policy, regional problems and arrangements, constraints that shape Presidential and departmental budgets, the interplay between foreign and domestic policy, political assessment and forecasting, uses and misuses of history, resource allocations and constraints, executive-legislative relations, government-press relations, and civilian-military relations.

It is intended that participants will leave not only with a better understanding of certain critical issues of national and international security policy, but also with a keener appreciation of the reasons for wide differences in perception and position relating to such issues and an enhanced ability to devise and implement effective solutions to complex and multifaceted problems.

The program is tailored to meet specific needs which senior officers and officials have identified as lacking in their own preparation for assuming posts at the highest levels in the national security community. It is a total immersion experience with a typical day starting at 7:00 a.m. and ending late that night with preparations for the following day's classes. During the course of the day, the participants are exposed to three 90-minute classes, and an hour of group discussion. There are a total of 30 such class sessions during the program, taught by leading experts in their fields. (See Attachments B and C for typical listings of assignments and faculty, respectively.) To facilitate informal interaction among participants, a number of social and athletic events are also scheduled.

Since the notion of perspective is such a critical ingredient of program substance, considerable effort is expended in achieving balanced representation from all segments of the national security community. See Attachment D for a typical roster of program participants.

ASSESSMENT

The analytical approach to the assessment of program effectiveness has remained essentially the same over the first three years of the program, comparing respectively the responses of military vs. civilian and War College vs. non-War College participants. Highlighted in this vein is an in-depth analysis of the features and patterns of gain in each participant's perceived competence in dealing with the various dimensions of national and international security. Cluster and correlation analysis techniques were introduced for this purpose during the second year of the program. This approach has been carried forward this year as well.

A comparison of the standard parameters of program assessment for the three years is presented in Table I.

TABLE I

A. On a scale of 1 (ineffective) to 5 (extremely effective):

	1978	1979	1980
1. Overall usefulness	4.3	4.6	4.4
2. Instructors	4.1	4.3	4.1
3. Cases	4.0	4.1	4.0
4. Administration	4.8	4.9	4.7

B. With respect to:

Program expectations

Exceeded	8	26	28
Equaled	20	18	41
Fell Short	1	0	4

Work load			
Too heavy	1	10	23
About right	29	34	51
Too light	1	0	0
Program length			
Too long	1	1	3
About right	29	36	64
Too short	1	7	5

In terms of achieving the principal objective of increasing student understanding of the perspectives of other participants in the national security process, an average grade of 4.5 was assigned during the third year. This compares with 4.2 and 4.4 respectively for the two previous years. Each year this score was refined through a "ratio scaling" technique wherein each executive was asked to evaluate on a scale of 100 to 1000 the degree to which he or she felt competent to deal with issues in each of the principal areas of interest (see Attachment E). In this regard, a series of three data points were taken: one prior to the program, one immediately following completion of the program, and one six months later (after sufficient time had elapsed to permit meaningful application of what was learned). In view of the heavily subjective nature of the total evaluation scheme, however, the pre-program data was taken at the same time as that immediately following the program. This was done to help ensure greater consistency in standards, i.e., the participants may not have fully appreciated at the outset how weak (or strong) they were in a given area as they were after they had completed the program. It also had the added benefit of avoiding an artificial "ceiling effect" wherein one might have given oneself the highest rating possible at the beginning of the program and, after learning considerably more during the program, then been forced to use the same (but subsequently misleading) rating.

While the method for assessing program effectiveness is highly subjective in nature, such an approach is dictated by the sensitivities involved in dealing with participants at this level. Their professional status is such as to preclude the possibility of canvassing their superiors for "before and after" impressions of performance. Attempts at evaluating classroom performance (including videotape analysis) proved equally futile since individual participation in a class size of 40 is too intermittent to provide a reliable indication. Another aspect complicating an objective evaluation of participant performance is the duality of the learning process. As Ernest May, Faculty Chairman for the Program, has noted, "the faculty learn as much from the experience as the students."

A trait common to most high level officials in evaluating the qualitative aspects of any training program is that of scrupulous honesty. For one thing, they feel an inherent sense of responsibility toward those who attend succeeding programs to provide the most candid assessment possible, even when it comes to evaluating themselves. For example, it is not at all unusual for a participant to evaluate his or her pre-program understanding of a given area at a level of 300 (on the 100 to 1000 scale).

As indicated in Table II, the average participant entered the program with a mean "competence" of 509 across all areas, i.e., the "average participant" felt about 51% as competent or comfortable in an "average area" as he or she did in an area with which they were thoroughly familiar and in which they felt well qualified (see Attachment F), During the course of the program, this mean index increased to 683, a 34% improvement over the pre-program level. Six months later, it had increased an additional 7% to 733. Comparable figures for the previous years are included in Table II and displayed in Figures I and II.

TABLE II
Average "Competence" Across All Areas

	Pre-Program	Post-Program	Six Months Post
Total Group (N=29) 1978	619	714 (14%)	741 (4%)
Total Group (N=30) 1979	589	714 (22%)	752 (5%)
Total Group (N=41) 1980	509	683 (34%)	733 (7%)

There was a significant decrease in the pre-program figures over the first three years of the program. As Table II indicates, the pre-program level dropped from 619 in 1978 to 509 in 1980. However, the six months post-program levels were approximately equal for each year's respondents. This decrease in pre-program levels may be at least partially attributable to a trend towards increased specialization in the represented population. For example, there was greater representation from the technical research and development community in each succeeding year. The fact that all groups reached approximately the same post-program levels suggests that the program is proving a certain effectiveness in bringing all groups to the same level of competence, independent of their perceived pre-program levels.

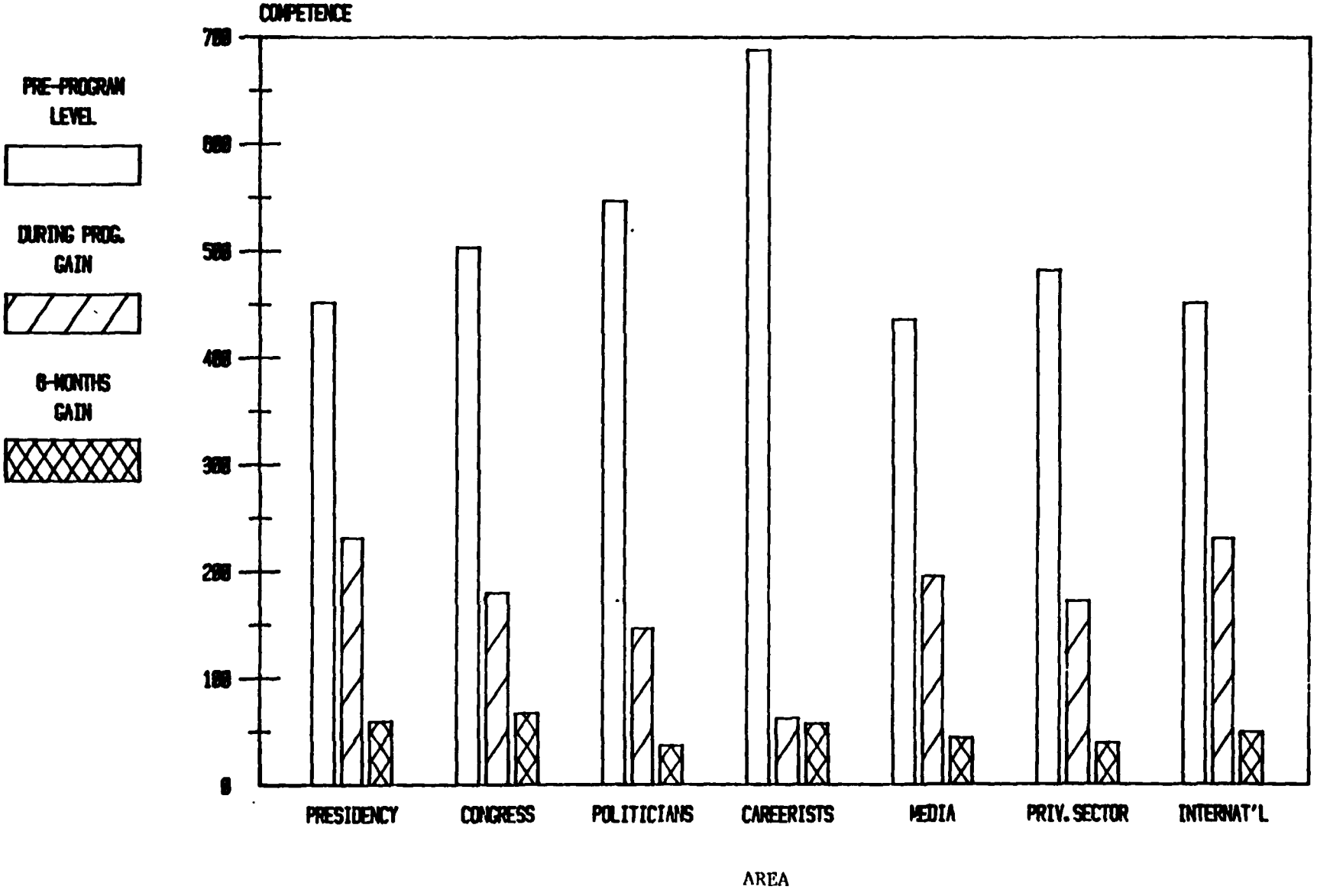
More revealing are the increases for each of the specific areas of emphasis. Table III and Figures IIIA and IIIB provide the indices for the 1980 participants in the seven areas of major interest.

TABLE III
Competence in Specific Areas (1980)

Area	Pre-Program Level	Gain During Program	Gain in Six Months
Presidency	451	232	58
Congress	502	180	67
Politicians	548	146	37
Careerists	688	62	57
Media	436	194	44
Private Sector	482	171	40
International	453	232	50

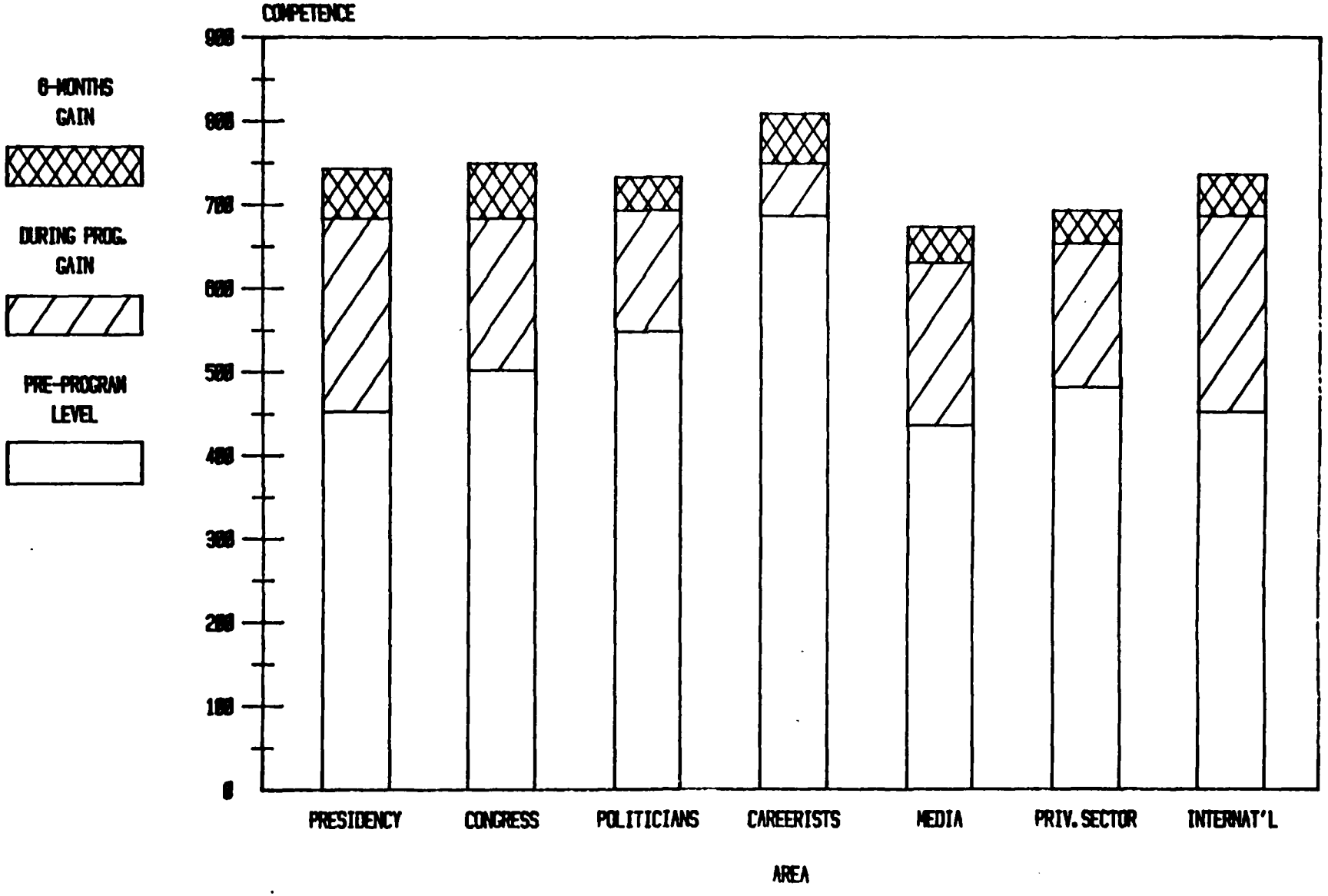
COMPETENCE IN SPECIFIC AREAS

FIGURE IIIA



COMPETENCE IN SPECIFIC AREAS (1980)

FIGURE IIIB



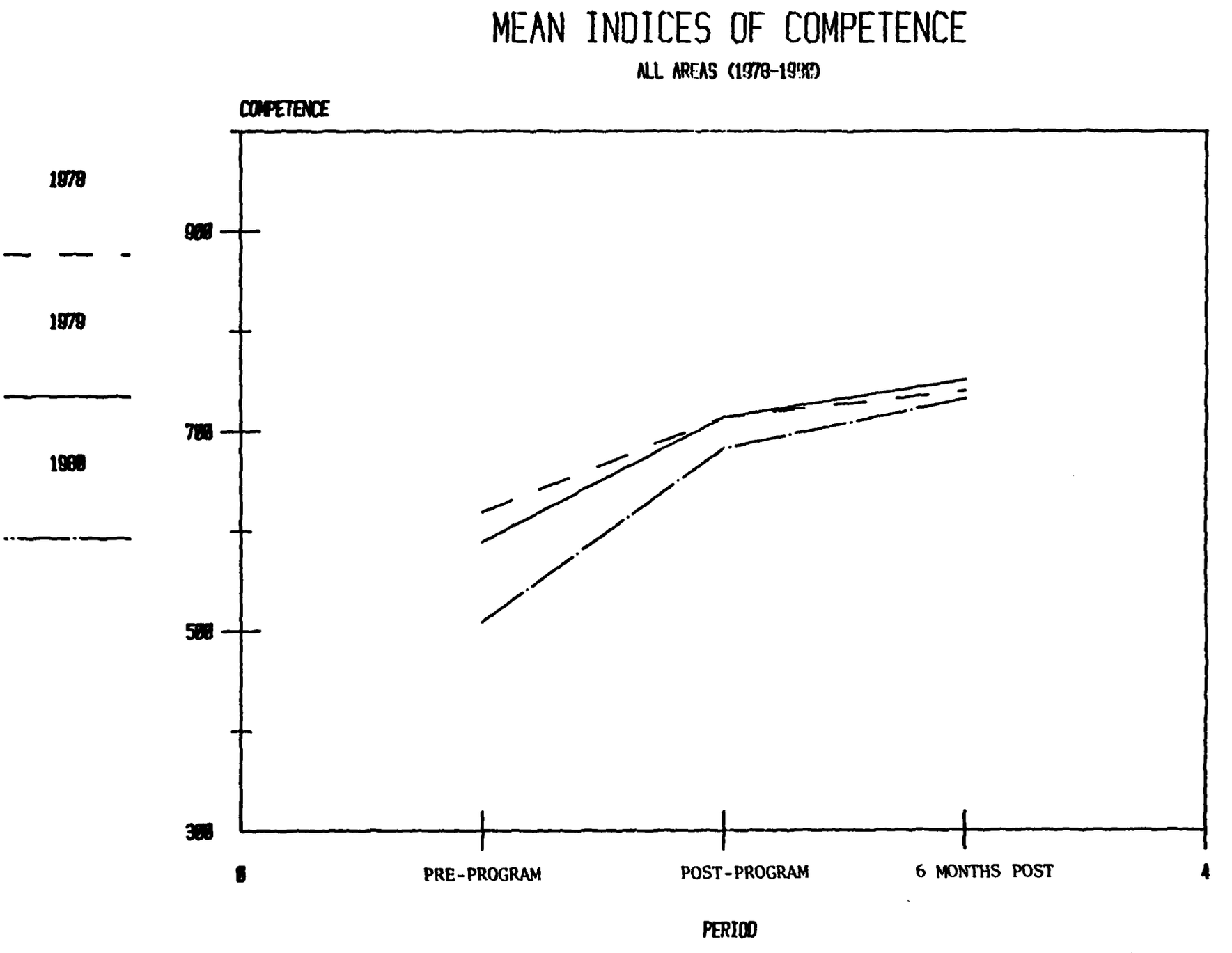

MEAN INDICES OF COMPETENCE
ALL AREAS (1978-1980)
COMPETENCE
1978
1979
1980
900
700
500
300
0
PRE-PROGRAM
POST-PROGRAM
6 MONTHS POST
4
PERIOD

FIGURE I

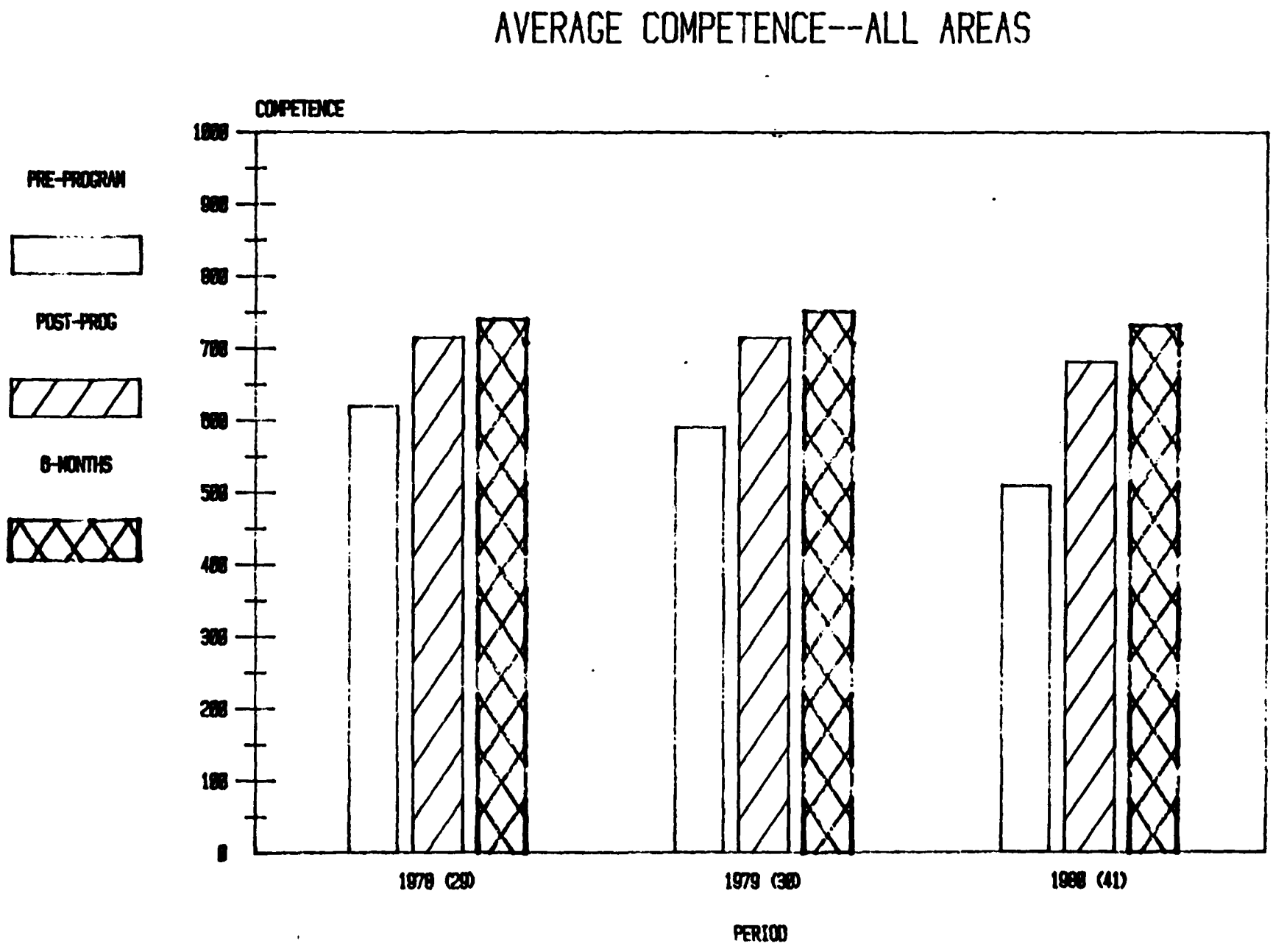

AVERAGE COMPETENCE--ALL AREAS
COMPETENCE
1000
900
800
700
600
500
400
300
200
100
0
PRE-PROGRAM
POST-PROG
6-MONTHS
1979 (29)
1979 (30)
1980 (41)
PERIOD

FIGURE II

Figure IIIA shows that the 1980 participants felt more competent in dealing with the "careerist" perspective that any other. This comes as no surprise since the majority of the participants are either military or civilian careerists themselves. As might be expected, they experienced little growth in this area as a result of the program. This is symptomatic of a broader pattern: the higher the pre-program competence, the less the gain during the program. Figure IIIB indicates that by the end of the program, all competence levels were close to equal; with the largest disparity occurring between the Media and Careerist categories.

This pattern is not unique to the 1980 group. Table IV provides a comparison of the gains in each of the categories common to all three years (the areas of emphasis changed slightly during the second year; thus there is no 1978 equivalent of the Presidency and International categories). Again, as might be expected, the high pre-program levels for Careerists, contrasts with those of the lowest, Private Sector and Media. By the six months post-program point, all areas were approximately equal. Figures IVA, B, and C further illustrate this point.

TABLE IV
Sample Group Comparisons

	Pre-Program			Program Gain			Six Months Post-Program Gain		
	1978	1979	1980	1978	1979	1980	1978	1979	1980
Careerists	699	781	688	46	38	62	22	41	57
Congress	615	591	502	105	113	180	38	54	67
Politicians	612	624	548	111	96	146	30	44	37
Private Sector	601	495	482	96	153	171	32	16	40
Media	589	536	436	88	154	194	28	46	44

It is noteworthy that there are no glaring differences among the three groups. Aside from the general downward trend in pre-program scores,

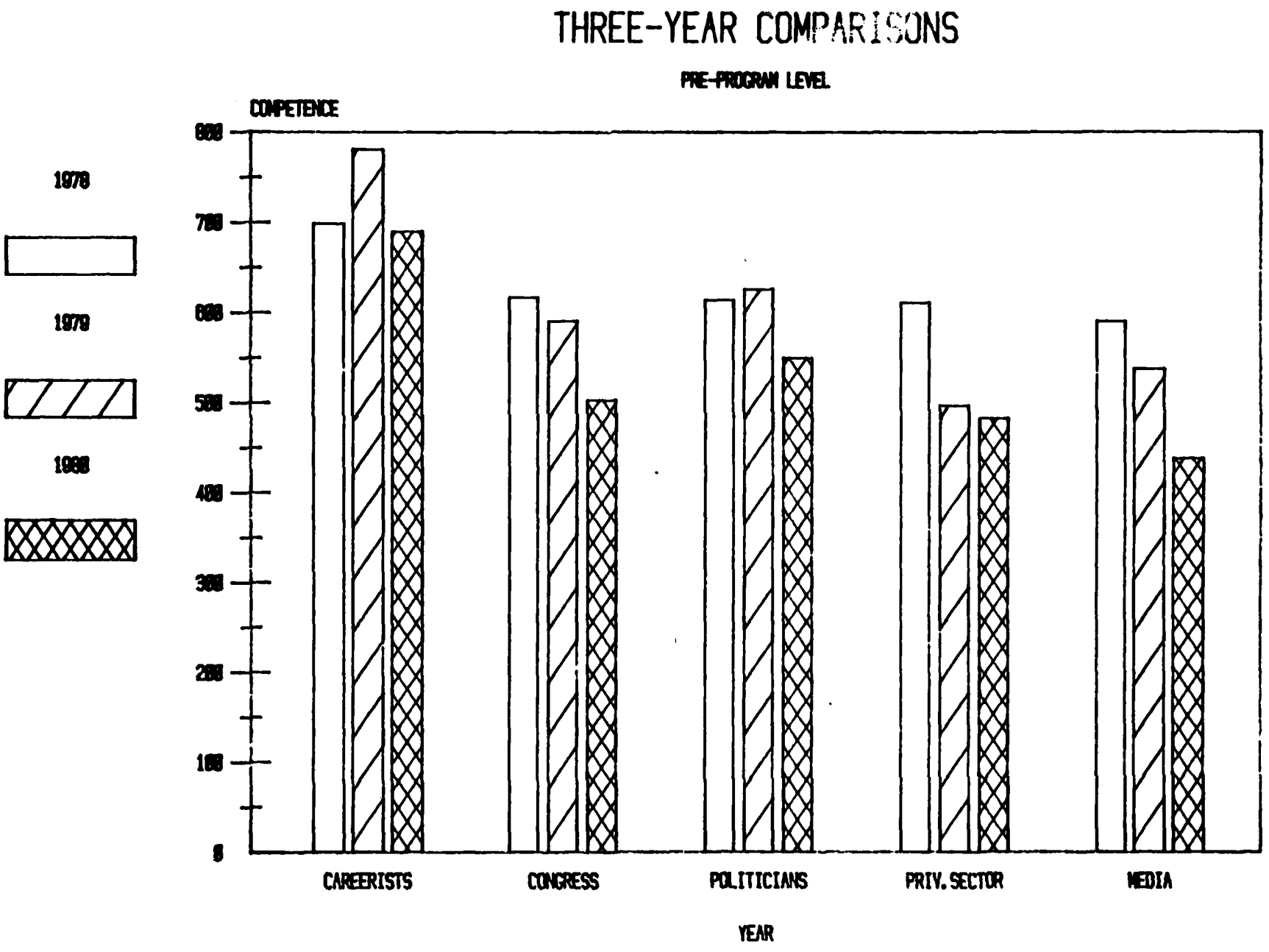

THREE-YEAR COMPARISONS
PRE-PROGRAM LEVEL
COMPETENCE
1978
1979
1980
CAREERISTS
CONGRESS
POLITICIANS
PRIV. SECTOR
MEDIA
YEAR

FIGURE IVA

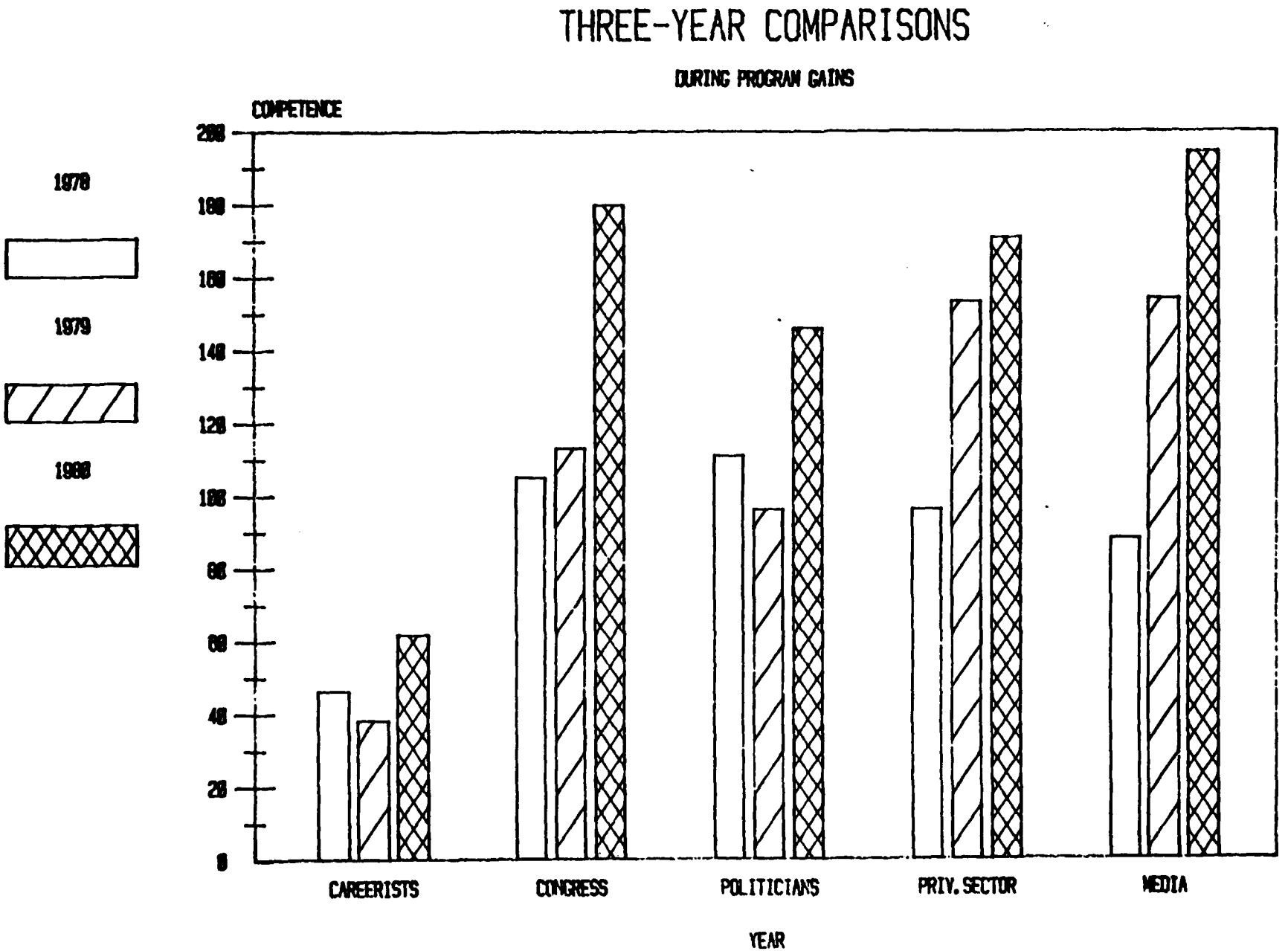

THREE-YEAR COMPARISONS
DURING PROGRAM GAINS
COMPETENCE
1978
1979
1980
200
180
160
140
120
100
80
60
40
20
0
CAREERISTS
CONGRESS
POLITICIANS
PRIV. SECTOR
MEDIA
YEAR

FIGURE IVB

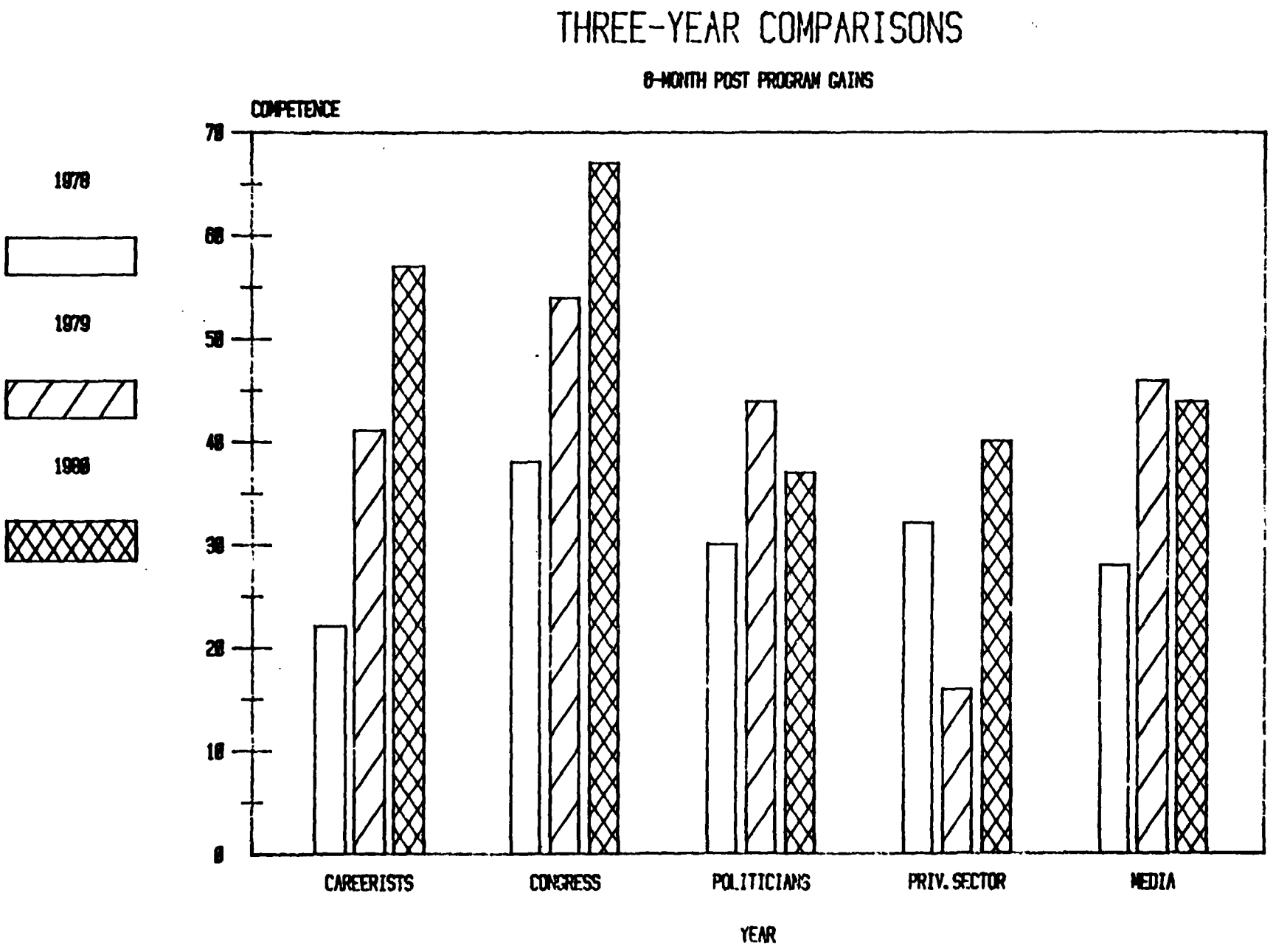

THREE-YEAR COMPARISONS
6-MONTH POST PROGRAM GAINS
COMPETENCE
70
60
50
40
30
20
10
0
1978
1979
1980
CAREERISTS
CONGRESS
POLITICIANS
PRIV. SECTOR
MEDIA
YEAR

FIGURE IVC

ultimate levels achieved in each of the five common areas has not varied much. Hence, there is no reason to believe that the three groups were not comparable in their reactions to the program, nor that the conclusions based on the responses from one group are not generally applicable to others.

Returning to Table IV, we specifically note that the average 1980 pre-program level is about 80 less than the corresponding 1979 level. The only area which is not significantly lower in 1980 is Private Sector, probably because it is the single area that is most uniformly foreign to government participants. Except for the areas of Private Sector and Careerists, the average "during program" gain is approximately 50 points higher in 1980 than in 1979. The reason for lesser Private Sector gain is probably due to the fact that the pre-program levels were approximately the same to begin with. The lack of growth in Careerists is probably due to a ceiling effect.

In comparing the reactions of those participants in the 1980 program who had previously attended War College with those who had not, the former entered the program feeling less competent (468 vs. 531 as opposed to the 1979 and 1978 experience of 647 vs. 594 and 647 vs. 594 respectively)* and left feeling that they had gained more (212 vs. 153 as was also true in prior years: 128 vs. 127 for 1979 and 103 vs. 91 in 1978). Figure V

---

* Why the War College participants evaluated themselves at a lower average pre-program level than their non-War College colleagues during the third year defies precise explanation. It may have to do with the fact that program attendance almost doubled between the second and third years. Since this expansion involved the acceptance of a number of less-experienced Flag officers (of which the typical War College class is mostly comprised), this may have had a dampening effect on the self-perceived pre-program index of competence for the War College group (and the military group as reflected in Figure VI).

illustrates this comparison for the last two years of program data. Thus it appears that the War College graduates eventually obtained a marginally higher level throughout than did their non-War College peers. This average level gain does not reflect the true effect of the program on these two groups, however.

A t-test for differences between means was applied to the War College and non-War College group averages in Table VA to determine whether the differences were statistically significant (Attachment G contains a discussion of the use of t-tests for this analysis).

TABLE VA
War College vs. Non-War College (1980)

	Pre-Program		Post Program		Six Months	
	War Coll	Non-War	War Coll	Non-War	War Coll	Non-War
Presidency	449	452	694	677	764	729
Congress	473	520	683	683	780	733
Politicians	433 *	613	678	702	763	712
Careerists	604 **	737	703	777	805	808
Media	422	444	647	621	703	658
Private	463	492	665	646	732	671
International	435	463	687	685	767	717
Average	468	531	680	684	759	718

War College = 15 respondents
Non-War College = 26 respondents

* significantly different @ 5% level
** significantly different @ 10% level

A related phenomenon of the expansion in class size was a perceptible diminishment in camaraderie. This was reflected in a noticeable decrease in the number of participants who responded to the six month post program questionnaire.

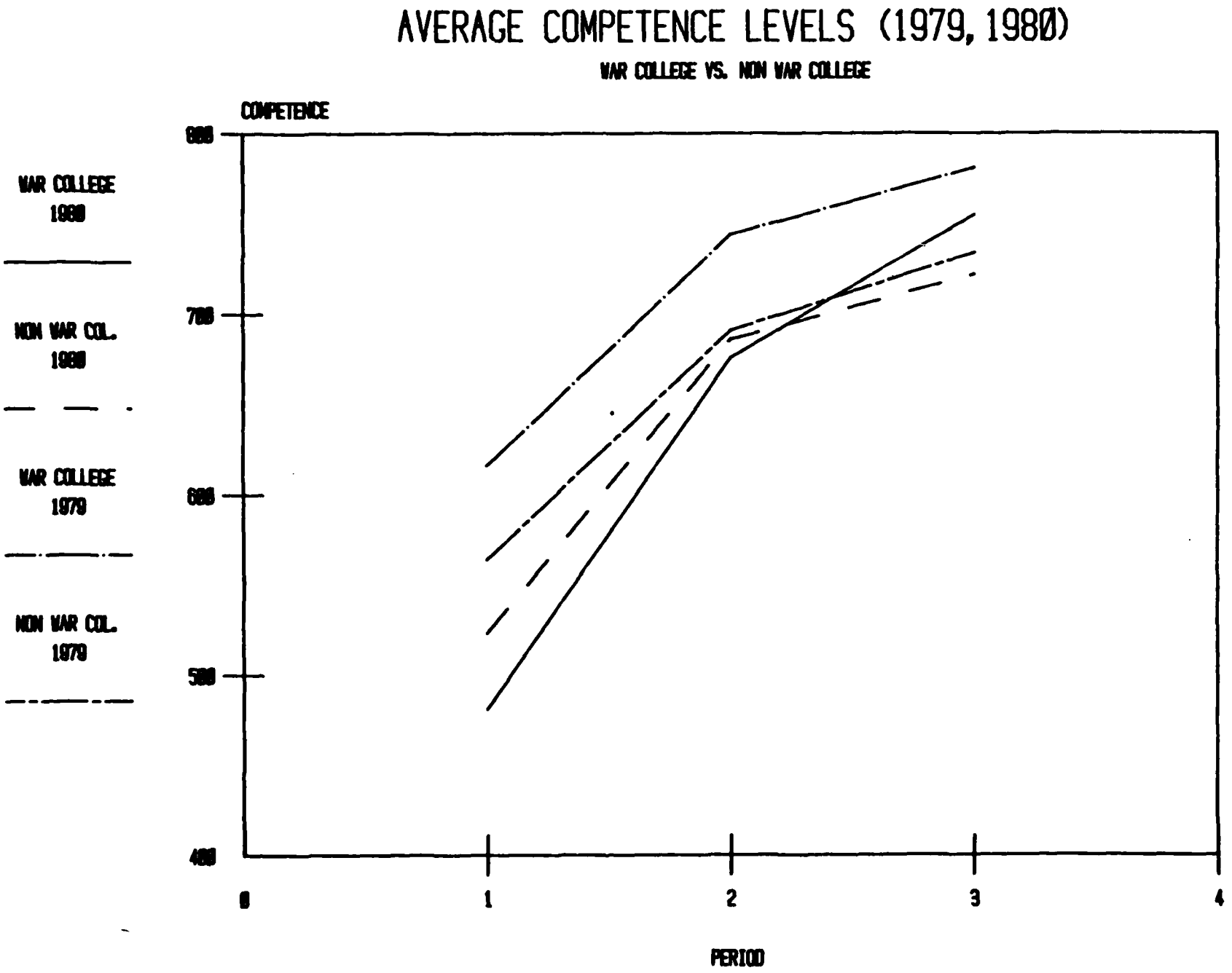

AVERAGE COMPETENCE LEVELS (1979, 1980)
WAR COLLEGE VS. NON WAR COLLEGE
COMPETENCE
WAR COLLEGE 1980
NON WAR COL. 1980
WAR COLLEGE 1979
NON WAR COL. 1979
0
1
2
3
4
PERIOD

FIGURE V

The only area in which the pre-program difference proved significant was that of Politicians, wherein War College graduates rated themselves significantly lower than did non-War College attendees (433 vs. 613). The two means, 433 and 613, are significantly different at a 95% confidence level. Using the equations in Attachment F:

	War College (WC)	Non-War College (NWC)
$\bar{x}$	433.3	613.5
S	247	207
N	15	26

The pooled variance $(S^2) = \dfrac{(N_{wc}-1)S^2_{wc} + (N_{nwc}-1)S^2_{nwc}}{(N_{wc}-1) + (N_{nwc}-1)}$

$= 49{,}103.$

The sample variance for differences of means $(S^2_{\bar{d}}) = \dfrac{S^2}{N_{wc}} + \dfrac{S^2}{N_{nwc}}$

$= 5{,}162.$

The t-statistic $(t_{\bar{d}}) = \dfrac{\bar{x}_{wc} - \bar{x}_{nwc}}{S_{\bar{d}}}$

$= 2.51$ with degrees of freedom 39.

t = 2.51 is significant at a .016 level (which means that two random samples from the same normal distribution will have a t-statistic as large as 2.51 only 1.6% of the time). Note: the standard deviations above are equal; that is, $S_{wc} \neq S_{nwc}$ is rejected at the 5% level.

Table VB makes a similar comparison, except the respondents are separated according to "military" or "civilian" status. Once again, the only area which demonstrated a 5% significant pre-program difference was Politicians.

TABLE VB
Competence in Specific Areas (1980)
Military vs. Civilian

Area	Pre-Program Level		Post-Program Level		Six Months Post Level		
	MIL	CIV	MIL	CIV	MIL	CIV	
Presidency	430	458	651	695	727	747	
Congress	481	510	685	682	777	740	
Politicians	436*	588	643	712	718	735	
Careerists	669	695	726	758	830	798	
Media	503	412	727**	595	745**	648	
Private Sector	495	477	707	633	766**	667	
International	434	460	705	678	741	733	
Average	493	514	692	679	758	724	198
Average	573	610	717	709	752	753	197

Military = 11 respondents
Civilian = 30 respondents

* significantly different @ 5% level
** significantly different @ 10% level

As can be seen in Table VI, the non-War College respondents gained significantly less (in a statistical sense) than War College graduates in the Politicians area only.

TABLE VI
During Program Gain
War College vs. Non-War College

Presidency	220	239
Congress	196	172
Politicians	209	113
Careerists	77	54
Media	234	174
Private Sector	198	157
International	230	233

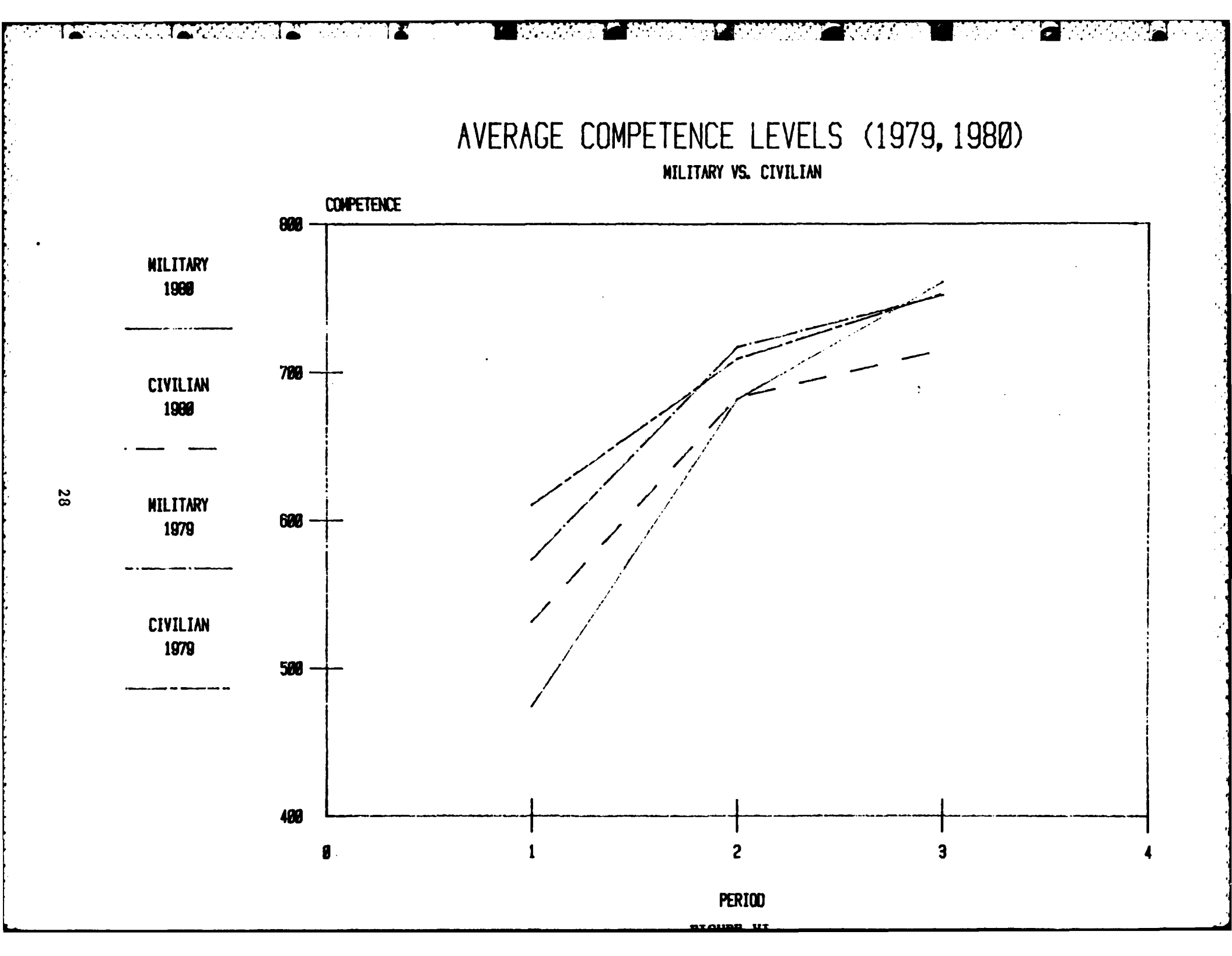
There is no apparent significance to this difference except to note that the only area in which the military respondents also significantly outgained their civilian counterparts was that of Politicians as well.

TABLE VII

	Military	vs.	Civilian
Presidency	249		222
Congress	221		154
Politicians	229		92
Careerists	93		42
Media	217		180
Private Sector	198		154
International	248		222

This is counter to the previous year's experience in which the non-War College group outgained the War College group in all areas except Media, Private Sector and International. For the record, in 1979, the military outgained the civilians in all areas. Figure VI compares the 1980 results with those for 1979.

Cluster analysis (5,6) was applied to the 1980 response data to determine if any specific identifiable subset of attendees benefitted significantly more than others who attended. This technique was initially applied last year to determine which participants had similar gain patterns in each of the seven areas of emphasis. Once the subsets of similar scoring participants were identified (these subsets are called "clusters"), then the individuals' backgrounds (career area, service, experience, etc.) were compared to determine whether there were any significant commonalities. These commonalities must, of course, be unique to the cluster so as to set it apart from other clusters. If such commonalities are found to exist, then it may be inferred that, inadvertently or not, the program favors a certain class of attendee. Thus, cluster analysis indicates among which individuals a common classification should be sought.


AVERAGE COMPETENCE LEVELS (1979, 1980)
MILITARY VS. CIVILIAN
COMPETENCE
800
700
600
500
400
0
1
2
3
4
PERIOD
MILITARY 1980
CIVILIAN 1980
MILITARY 1979
CIVILIAN 1979

Summarizing the results of the cluster analysis of the 1979 data: two small tightly packed clusters and three distinct "outliers" were discerned. Comparing the backgrounds of the participants in the two clusters and the three outliers yielded the following:

1. There were no significant elements of commonality in the two clusters.
2. The three "outliers" had backgrounds similar to many who were much closer to the clusters.
3. The average distance from case to case among all the cases was sufficiently large to conclude that the cluster analysis had not highlighted an identifiable subgroup of attendees who drew inordinate benefit (or obtained no benefit at all) from the program.

The same analytical technique (in fact, the same computer program) was applied to the "during program" gain data for the 41 respondents from the 1980 program with nearly identical results (see Attachment H for cluster analysis algorithm and computer analysis). Instead of the two tightly packed clusters, there were two more loosely associated groups, but again, with widely varying backgrounds. There were six outliers, half of whom were military personnel, with two of them having had War College experience. It was therefore concluded (even more strongly than it was in 1979) that the executive program's method and curriculum are not biased towards or away from any specific, identifiable subset of participants.

As explained in last year's report, correlation analysis (4,5) is used to quantify the degree of association among scores, i.e., if an attendee gains a great deal during the program in area A, can he or she be expected to gain significantly in area B? For clarification: the charts and tables comparing mean competence levels and mean gain levels among different years

and groups indicate how well different participants have responded to the program curriculum. The cluster analysis indicates whether a particular subset of attendees benefits more from (or responds better to) the program curriculum; in other words, it addresses the program design. Correlation analysis, on the other hand, indicates how consistently individual participants respond to the seven different areas of emphasis (Attachment I contains a mathematical description of the correlation statistics).

Recalling the results of the correlation analysis on the 1979 data: the scores in the Presidency category served as a good proxy for all gain scores for an individual; that is, the program gains for Presidency correlated highly with all other areas. However, there did not appear to be any synergistic effects, in which training in one area was seen as perfectly complementing that in another such that there were inordinate gains in both, or that training in one area supplemented that in another so well that equal gains would be realized in both (such effects would have been made apparent by singularly high correlations in one or two cells of the matrix, with relatively low levels elsewhere). In general, there appeared to be an overall complementary effect to the training.

The analysis has been extended for the 1980 data. Correlation statistics have been computed and compared for both pre-program levels and during program gains for the military vs. civilian and War College vs. non-War College groups. Tables VIII a-h present the individual correlation matrices.

Looking at the civilian vs. military comparison and focusing on the pre-program levels (Tables a and c), it is obvious that the military attendees' responses correlate much higher in all areas (with the exception of the "Congress") than do their civilian colleagues, and lower in very few.

CORRELATION MATRIX

TABLE VIII A

CIVILIAN : PRE PROGRAM LEVEL

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.3	1.0					
POLI	.0	.3	1.0				
CARE	- .2	.1	.6	1.0			
MEDI	.3	.1	.2	.1	1.0		
PRIV	.3	.3	.3	.5	.4	1.0	
INTR	.4	.4	.2	.3	.5	.5	1.0

CORRELATION MATRIX

TABLE VIII B

CIVILIAN : DURING PROGRAM GAIN

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.3	1.0					
POLI	.4	.1	1.0				
CARE	.4	.1	.8	1.0			
MEDI	.2	.0	.3	.3	1.0		
PRIV	.4	.2	.1	.3	.3	1.0	
INTR	.3	.4	.2	.3	.4	.5	1.0

CORRELATION MATRIX

TABLE VIII C

MILITARY : PRE PROGRAM LEVEL

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.8	1.0					
POLI	.5	.1	1.0				
CARE	.6	.6	.4	1.0			
MEDI	.4	.4	.4	.6	1.0		
PRIV	.7	.8	.8	.3	.3	1.0	
INTR	.2	.4	.4	.5	.4	.3	1.0

CORRELATION MATRIX

TABLE VIII D

MILITARY : DURING PROGRAM GAIN

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.5	1.0					
POLI	.4	.6	1.0				
CARE	.3	.0	.4	1.0			
MEDI	.2	.0	.2	.4	1.0		
PRIV	.5	.6	.5	.1	.1	1.0	
INTR	.4	.6	.3	.1	.1	.6	1.0

CORRELATION MATRIX

TABLE VIII E

NON WAR COLLEGE : PRE PROGRAM LEVEL

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.2	1.0					
POLI	- .3	.3	1.0				
CARE	- .3	.3	.6	1.0			
MEDI	.1	.1	.0	- .1	1.0		
PRIV	.3	.4	.1	.4	.3	1.0	
INTR	.2	.6	.0	.2	.4	.4	1.0

CORRELATION MATRIX

TABLE VIII F

NON WAR COLLEGE : DURING PROGRAM GAIN

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.4	1.0					
POLI	.2	.0	1.0				
CARE	.1	.2	.7	1.0			
MEDI	.2	.1	.1	.1	1.0		
PRIV	.3	.5	.0	.1	.3	1.0	
INTR	.2	.6	.0	.2	.4	.6	1.0

CORRELATION MATRIX

TABLE VIII G

WAR COLLEGE : PRE PROGRAM LEVEL

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.7	1.0					
POLI	.7	.6	1.0				
CARE	.4	.1	.3	1.0			
MEDI	.5	.2	.4	.6	1.0		
PRIV	.6	.4	.7	.5	.5	1.0	
INTR	.5	.1	.5	.6	.6	.5	1.0

CORRELATION MATRIX

TABLE VIII H

WAR COLLEGE : DURING PROGRAM GAIN

	PRES	CONG	POLI	CARE	MEDI	PRIV	INTR
PRES	1.0						
CONG	.1	1.0					
POLI	.7	.4	1.0				
CARE	.6	- .1	.6	1.0			
MEDI	.4	- .4	.5	.5	1.0		
PRIV	.4	.0	.3	.2	.2	1.0	
INTR	.6	.1	.5	.3	.3	.4	1.0

Tables e and g provide the same comparison for the non-War College vs. War College with the latter category correlating much higher than the former. These higher correlations indicate that the military/War College attendees are more consistent in their responses than are the civilian/non-War College participants. There are two phenomena at work here. One is that the military and War College experiences tend to impart similar biases and points of view to those involved. The other is that some military personnel, depending on their assignments, tend to consider the Congress off-limits, whereas civilians feel no such inhibition. Thus we see that the military and War College groups tend to be more single-minded. While they don't necessarily score higher than the others, they do score with greater consistency. There is a similar, but less marked, pattern appearing in the "during program" gain correlation matrices. The similar correlations between the military and War College groups are consistent with what one would expect since the vast majority of War College attendees are members of the Armed Forces.

EXPLORATION OF POSSIBLE METHODOLOGIES

During the first year of the program, three separate areas of study were subjected to mathematical or statistical analysis, each of which related to the overall objective of determining training effectiveness. The investigation subjects included:

1. Analysis of measures of effectiveness.

   The investigations in this area were directed toward assessment of participant responses to increase their perceptions of how the program had affected them. The analysis involved the use of simple descrip-

tive statistics (means, variances, etc.) for each of the program subject areas, trend analysis over the three years, determination of subpopulation scoring patterns, and the like. These investigations proved most fruitful and were continued in subsequent years.

2. Determination of long-term effects.

There was a clear trend of growth in all areas of competence both during the program and for the six month period following its completion. The dimensions addressed in this area of investigation included: (1) how long the program's influence would be felt by the participants, and (2) to what levels of competence they could be expected to rise. Attempts to model the growth pattern as a non-stationary stochastic process were not particularly successful. Additional data beyond that collected at the six month post-program point would have been required to yield conclusive results. Since it was felt that any further demands on participants' time in this regard would have represented an imposition, this area of investigation was discontinued.

3. Prediction of applicant population.

Another, less direct, measure of effectiveness was the degree to which the program was capable of attracting an increased number of applicants in successive years. A theoretical model of the application/attendance process was developed. However, since the model was based on the then current methods of recruiting, which changed dramatically over the succeeding years, its utility as a predictive tool was judged insignificant.

Methodologies

1. Measures of Effectiveness.

The basic approach to analyzing measures of effectiveness was to identify descriptive statistics that characterized the patterns of competency scores (12). Other than the simple histograms and time plots shown in this report, statistical comparisons have been addressed among such groups as military vs. civilian, War College vs. non-War College, 1979 vs. 1980, and area vs. area with such analytic tools as t-tests, F-ratio tests, correlation analyses, and cluster analyses. There are, of course, many other analytical methodologies that could have been employed. Instead of the simple correlations, covariance analyses, regression analyses, and the like could have been performed. In addition to cluster analysis, factor analysis (3,5), logit and probit analysis (10), discriminant analysis (2,3), and principal components analysis (2,5) could have been performed. However, the correlation and cluster techniques provide intuitive results, even for the non-analyst. Most of the other techniques would have required manipulation and transformation of the data, often in a sense that would probably bury interpretive results. Note that a rigorous experimental design was not employed wherein it was decided ahead of time to analyze the effectiveness measure with principal component methodologies, for example, and then construct a questionnaire to provide the required parameters of such a model.

The questionnaire that was used was designed to give maximum information on the respondents' opinions of the program and its effectiveness and on their associated personal growth in competency in specific

areas of national security concern. Preliminary analysis of these data indicated that there may have been some underlying structure to the participant scores. Hence the comparisons of military vs. civilian, etc. Note, again, most of the methodologies mentioned above have multiple independent variable and single dependent variable models. We did not have such a structure to work with. Instead, the individual scores had to be considered as dependent variables, while civilian or military or War College experience, etc., was the sole independent variable available. [Some factor analysis techniques could have been applied, using military/civilian as one independent variable and War College/non-War College as another; however, none of the models can provide meaningful results when there are more dependent than independent variables. Hence one would have had to select at most two scores as dependent variables or build some sort of composite dependent variable for each participant. It is such manipulation and transformation mentioned above that would fog the issues -- it just doesn't make good mathematical/engineering sense to contort the data to fit an analytical model (12).] This is the reason descriptive statistics were used instead of some of the higher-powered statistical models. However, cluster analysis was conducted to determine if there existed a real need to perform such analyses. Recall that in the cluster analysis, one was looking for subsets of the population that scored very similar to one another, but different from the rest of the population. If such subgoups had existed, this would have provided good indication that there were, indeed, nonobserved independent variables that were driving the scoring patterns. The results of the cluster analysis made it quite clear that there were no such sub-

groups, and therefore, no independent variables that were crucial to the understanding of the scoring patterns. [Such independent variables could have been age, income, educational background, military service branch, sex, geographical location of domicile and birth, etc.]

In summary, it was found that descriptive statistics provided sufficient visibility into "who scored what" and that there was no statistical basis for determining "why."

2. Long-term Effects.

There exists a library-full of stochastic models that could be used to attempt to model the long-term effect of the program training on the participants. In fact, in the first year's report, the attempt was made to fit a stationary (and then a non-stationary) Markov model to the growth in competence. Data was first manipulated and transformed from raw scores to a simple set of high, medium, and low scores. The transition probabilities that an individual respondent would move from one category (high, medium, or low) to another over the program and then in six month intervals were computed. While it was possible to model the first three sample points (pre-program, post-program, and six months post-program), and even the fourth (12 month post-program) with a sufficiently complex model, the model was of little, if any, predictive value. From the data, it was impossible to determine when the growth process would reach a steady state, and even what the nature of that steady state would be. In effect, the problem was to choose the "best fit" to three points; however, this fit did not

provide a good approximation of the fourth point. The basic problem was one of attempting to extrapolate approximately 100% beyond the total time span of the data sample. This is a frequent downfall of time-series modelling: forecasting beyond the sample. When there are long time span samples (long relative to the required forecast period) then time-series forecasting using Markov (12) models (as well as other models such as Box-Jenkins, ARIMA, semi-Markov, Power Spectrum, etc.) is reasonable. The samples in this case are just too short-term to provide reliable indication about the long-term. The analysis that was performed is attached for the reader's information.

3. Applicant Prediction.

In attempting to predict the future population of applicants to the program, it was postulated that the application process within each organization that was a potential supplier of candidates was the result of a Poisson process with an average application rate, λ, for each organization. Furthermore, λ was assumed to be a linear function of the then extant recruiting policies (number of brochures distributed, number of visits by the program's executive director, Service's budget for such training, etc.) Finally, it was hypothesized that the final pool of applicants resulted from a Gamma distribution. (Overall, this meant that the application process was modeled as a random negative binomial process.) The approach proved quite realistic with the first year's data fitting the model rather well (1,4).

Formally, the model was stated as follows:

A. $\lambda_i = \beta_o + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i}$

where λ_1 = # of applicants from agency i

x_{1i} = training budget for agency i

x_{2i} = # of eligible candidates for agency i

x_{3i} = # of brochures mailed to agency i

x_{4i} = # of visits to agency i

β = regression coefficients

B. Probability that (# attend from agency i = κ) = $\frac{\lambda_i^{\kappa} e^{-\lambda i}}{\kappa!}$

C. λ_i is distributed as a Gamma:

$$F(\lambda) = \begin{cases} \frac{1}{P(v)} a^v \lambda^{v-1} e^{-a\lambda} & \text{if } \lambda > 0 \\ 0 & \text{if } \lambda = 0 \end{cases}$$

D. Resulting in the Negative Binomial:

$$\text{Prob } (\kappa=n) = q^v p^n (-1)^n \binom{-v}{n}$$

where $p = \frac{1}{1+a}$ and $q = 1-p$

and mean $\mu = vp/q$ and variance $\sigma^2 = \mu(1+\mu/v)$

While the first year's data when used with this model, as well as a simple Poisson model, fit the sample distribution rather well, the second year's data did not fit the established model (A above) at all. Thus, the changes in recruiting approach from one year to the next preempted the model's usefulness as a predictor of future applicants.

SUMMARY

Summarizing the results of the analysis of the 1980 data, we conclude as we did in 1979 that, foremost, the executive program clearly raises the perceived level of competence among all who participate, independent of the participant's background (in terms of military vs. civilian and War College vs. non-War College). The trend over the first three years toward reduced pre-program levels of competence has been largely matched by an offsetting trend toward increased program gains, indicating a certain degree of program effectiveness in raising all groups to a similar level of competence. It has been further demonstrated that the separate elements of the program tend to complement one another. While it has also been shown that the program curriculum is not directed toward any specific group of participants, the military and war college attendees have characteristically responded in a more consistent manner to the full spectrum of program training.

BIBLIOGRAPHY

1. Bishop, Y., S. Fienberg, P. Holland, Discrete Multivariate Analysis, MIT Press, Cambridge, 1975.

2. Cooley, W.W. and P.R. Lohnes, Multivariate Data Analysis, John Wiley & Sons, New York, 1971.

3. Fienberg, S.E., The Analysis of Cross-Classified Categorical Data, MIT Press, Cambridge, 1977.

4. Fisz, M., Probability Theory and Mathematical Statistics, John Wiley & Sons, (3rd edition), New York, 1963.

5. Gnanadesikan, R., Methods for Statistical Data Analysis of Multivariate Observations, John Wiley & Sons, New York, 1977.

6. Hartigan, J.A., Clustering Algorithms, John Wiley & Sons, New York, 1975.

7. Kashyap, R.L., Dynamic Stochastic Models from Empirical Data, Academic Press, New York, 1976.

8. National Academy of Public Administration, "Strengthening Civilian Executive Development in the Department of Defense," March 1976.

9. Navy Personnel Research and Development Center, "Navy Civilian Executive Study," January 1979.

10. Nerlove, M. and S.J. Press, Univariate and Multivariate Log-Linear and Logistic Models, RAND Report R-1306-EDA/NH, 1973.

11. Pask, G. and M. Robinson, "Current Scientific Approaches to Decision Making in Computer Systems: III," Volumes 1 and 2 of report on conference held at Richmond, Surrey, England, under the auspices of the U.S. Army Research Institute for the Behavioral and Social Sciences and European Research Office, U.S. Army, London, August 6-8, 1978.

12. Tufte, E.R., Data Analysis for Politics and Policy, Prentice-Hall, Inc., Inglewood Cliffs, New Jersey, 1974.

13. U.S. Army Management Engineering Training Agency (AMETA), "Study of Management and Executive Development in Industry, Universities, and the Federal Government," May 1974.

Attachment A

NEW CASE DEVELOPMENT FOR EXECUTIVE PROGRAM

	Title	Faculty
1.	Dienbienphu	May
2.	Bay of Pigs	May
3.	Panama Canal Treaty Negotiations (A) and (B)	Raiffa/Johnston
4.	United States - Phillipines Military Base Negotiations	Raiffa/Johnston
5.	Eurocommunism (A) and (B)	Yergin/Johnston
6.	Manpower Management in the Navy (A) and (B)	Johnston
7.	Middle East Negotiations: The Camp David Summit	Raiffa
8.	Congress and the F16	Moore/Zimmerman
9.	Energy Exercise	Yergin
10.	The Defense Budget	Lynn
11.	East Asian Security	Nacht
12.	European Security	Nacht
13.	AWACS (A) and (B)	Johnston
14.	Defense Management	Lynn
15.	The Intelligence Process: Operation Barbarossa	May
16.	Political Instability	Nacht
17.	The Cruise Missile and NATO Standardization	Johnston
18.	The President, the Congress, and the Nuclear Carrier (A) and (B)	Johnston

SCHEDULE OF CLASSES

DAY	PERIOD	SECT A (Room 150)	SECT B (Room 130)
(Aug 11)	1	Bay of Pigs/Cuban Crisis (May)	- Combined Session
	2	Decision Analysis (Raiffa)	2 C5A (Lynn)
	3	C5A (Lynn)	3 Decision Analysis (Raiffa)
Aug 12	1	Dienbienphu & Pleiku (May)	1 Defense Budget (Lynn)
	2	Bargaining Theory (Raiffa)	2 Dienbienphu & Pleiku (May)
	3	Defense Budget (Lynn)	3 Bargaining Theory (Raiffa)
Aug 13	1	International Economics I (Uyterhoeven)	- Combined Session
	2	Middle East (May/Nacht)	- Combined Session
	3	International Economics II (Uyterhoeven)	- Combined Session
	5:00 - 6:30 p.m.	Panel Discussion: The Policymaker and Intelligence (Rm 140)	
Aug 14	1	Congress and the F16 (Moore)	1 SALT I/II (Carnesale)
	2	Media I (Miller)	- Combined Session(Rm 140)
	3	SALT I/II (Carnesale)	3 Congress and the F16 (Moore)
Aug 15	1	Organization Analysis (Porter)	1 International Negotiations (Raiffa)
	2	Media II (Miller)	- Combined Session
	3	International Negotiations (Raiffa)	3 Organization Analysis (Porter)
Aug 16	1	Food (McCraw)	- Combined Session
	2	Grain (Porter)	- Combined Session

		SECT A (Room 130)	SECT B (Room 150)
(Aug 18)	1	Marxism/Communism (May)	- Combined Session
	2	Strategic Forces & Doctrine (Carnesale)	2 E. Asian Security (Nacht)
	3	E. Asian Security (Nacht)	3 Strategic Forces & Doctrine (Carnesale)
Aug 19	1	Proliferation (Carnesale)	- Combined Session
	2	Nuclear Exports (Nye)	2 Energy I (Uyterhoeven)
	3	Energy I (Uyterhoeven)	3 Nuclear Exports (Nye)
	5:00 - 6:30 p.m.	Panel Discussion: Executive/Legislative Relations (Rm 150)	
Aug 20	1	Energy II (Uyterhoeven)	1 Force Planning (Johnston)
	2	Energy Security (Nye)	2 Energy II (Uyterhoeven)
	3	Force Planning (Johnston)	3 Energy Security (Nye)
Aug 21	1	International Economics III(Uyterhoeven)	- Combined Session
	2	Dealing with Conflict (Fisher)	2 European Security (Nacht)
	3	European Security (Nacht)	3 Dealing with Conflict (Fisher)
Aug 22	1	Energy Exercise (Nye)	- Combined Session
	2	Wrapup (Johnston)	- Combined Session

Attachment C

FACULTY

Graham T. Allison (KSG)
(617) 495-1122

James E. Austin (HBS)
(617) 495-6597

Jonathan Brock
(206) 543-4920

Albert Carnesale (KSG)
(617) 495-1189

Antonia H. Chayes (KSG)
(617) 495-1149

John T. Dunlop (HBS)
(617) 495-1431

Roger Fisher (HLS)
(617) 495-4615

Stephen B. Hitchner
(202) 633-2661

Douglas M. Johnston, Jr. (KSG)
(617) 495-1331

William W. Kaufmann (MIT)
(617) 253-3376

Laurence E. Lynn, Jr. (KSG)
(617) 495-1323

Ernest R. May (KSG)
(617) 495-1144

Thomas K. McCraw (HBS)
(617) 495-6513

Arthur R. Miller (HLS)
(617) 4954111

Jonathan Moore (KSG)
(617) 495-1360

Michael L. Nacht (KSG)
(617) 495-1405

Joseph S. Nye, Jr. (KSG)
(617) 495-1148

Roger B. Porter
(202) 456-6722

Don K. Price (KSG)
(617) 495-1315

Howard Raiffa (KSG)
(617) 495-6289

Hugo E.R. Uyterhoeven (HBS)
(617) 495-6622

Daniel H. Yergin (KSG)
(617) 495-1359

Attachment D

EXECUTIVE PROGRAM
IN
NATIONAL AND
INTERNATIONAL SECURITY

August 16 - 28, 1981

John F. Kennedy
School of Government

HARVARD UNIVERSITY

ALESSI
Victor E.
Chief, Strategic Affairs Division
Arms Control & Disarmament Agency

AMATO
Joseph
Chief, Operations Group
Natl Security Agency

ARNOLD
Terrell E.
State Dept.
Advisor, National Defense University

ATKESON
Edward B.
MGEN, U.S. Army
Cdr, Army Concepts Analysis Agency

BENDER
Charles F.
Head, Chem. & Mats Science Dept.
Livermore Laboratory

BOATNER
Helene L.
Director of Political Analysis
Central Intell. Agency

BRADFORD
Zeb B., Jr.
BGEN, U.S. Army
Chief, Policy Branch, SHAPE

CLYNE
Norman G.
Executive Secy.
Arms Control & Disarmament Agency

DUNN
Robert F.
RADM, U.S. Navy
Commander, Navy Military Personnel Command

DURHAM
Archer L.
BGEN, U.S. Air Force
Cdr., 76th Air Lift Div., Military Air Lift Command

ECKMAN
Philip K.
Dir, Office of Research & Dev.
Central Intell. Agency

EWING
Ronald W.
Tech. Director, Foreign Tech. Div.
Air Force Syst. Cmd.

FANT
Joseph L.
MGEN, U.S. Army
Cdr, Army Readiness & Mob. Region I

FETTERMAN
John H., Jr.
RADM, U.S. Navy
Cdr., Tactical Wings U.S. Atlantic Fleet

FOX
Eugene
BGEN, U.S. Army
Dep. Dir., Defense Test & Evaluation
Offc of Secy of Def.

FRAGNER
Berwyn
MGEN, U.S. Army Reserve
Cdr, 63rd Army Reserve Command

JURKOWICH
George J.
Vice Pres for Public Rel. & Asst. to Pres.
Chase Manhattan Bank

KARBER
Phillip A.
Vice Pres., Natl Security Programs
BDM Corporation

KEELEY
John B.
Military Affairs Commentator
Cable News Network

KEISER
Richard E.
Dep. Asst. Director
Protective Research
U.S. Secret Service

KELLY
David Barry
Office Director, Science & Tech
Central Intell. Agency

KENNEDY
J. Keith
Staff Director
Appropriations Committee
U.S. Senate

KERN
Richard N.
Dep. Director, Plans & Policies
Nat'l Security Agency

KICKLIGHTER
Claude M.
BGEN, U.S. Army
Asst Chief of Staff Logistics, Allied Forces Cent. Europ

McGANN
Donald A.
BGEN, U.S. Air Force Reserve
Mobilization Asst to Cdr Aeronaut Syst Div

McHUGH
James J.
RADM, U.S. Navy
Dep. Judge Advocate General of the Navy

McKNIGHT
Clarence E., Jr.
MGEN, U.S. Army
Commanding General, U.S. Army Signal Center & Ft. Gordon

MEYERS
Robert S.
Dir., U.S. Foreign Press Ctrs, Internatl Communications Agency

NEILSEN
Finn K.
Deputy Assistant Administrator for Contingency Planning
Dept. of Energy

O'MALLEY
Edward J.
Asst. Director Intelligence Div.
Federal Bureau of Investigation

PEAT
Randall D.
BGEN, U.S. Air Force
Asst. Dept. Dir. Force Dev & Strat. Plans
Joint Chiefs of Staff

PERLSTEIN
David E.
Senior Team Leader, Program Assessment Office
Int. Community Staf

SILVERSTROM
Leon
Asst. Gen. Counsel for Internatl Dev. & Defense Progs.
Dept. of Energy

SITRIN
David
Deputy Assoc. Dir for Natl Security
Offc of Mgt & Budget

SMILEY
Annette H.
Prof. Staff
House Intelligence Committee

SNEDIKER
David K.
Mngr., Transportation & Structures Dept.
Battelle Laboratories

SOYSTER
Harry E.
BGEN (sel.), U.S. Army
Chief of Staff
24th Infantry Div.

TANKERSLEY
Will Hill
MGEN, U.S. Army Reserve
Mobilization Asst to DCS, Ops & Plans

TAYLOR
Edward C.
Director for Requirements Anal., Defense & Space Systems, TRW

THOMAS
Clayton J.
Scientific & Tech Advisor, Strategic Force Analysis
HQ, U.S. Air Force

Faculty

AUSTIN
James E.
Harvard Business School

CARNESALE
Albert
Kennedy School of Government

CHAYES
Antonia H.
Kennedy School of Government

FISHER
Roger
Harvard Law School

JOHNSTON
Douglas M., Jr.
Executive Director
Kennedy School of Government

LYNN
Laurence E., Jr.
Kennedy School of Government

MAY
Ernest R.
Faculty Chairman
Kennedy School of Government

COURTNEY
Charles E.
Princ. Sr. Inspector
International
Communications
Agency

CULPEPPER
James W.
Acting Principal Dep.
Asst. Secy for
Defense Programs
Dept. of Energy

DARLING
Robert E.
Asst. Dep. Chief
of Staff for Maint.
Air Force Logistics
Command

DAVIS
William A.
Deputy Program Mngr
Ballistic Missile
Defense
Dept. of Army

De ANGELUS
Alfred R.
Asst. Commissioner
Offc Commercial Ops
U.S. Customs Service
Dept. of Treasury

DESAULNIERS
Charles L.
Defense Intell.
Officer, E. Asia
& Pacific
Def. Intell. Agency

DUBICKI
Henry J.
Dep. for Acquisition
Policy, Offc of Secy
of Army

DUKE
John V.
Chief, Plans &
Policies Division
Def. Logistics Agency

FRICK
Roy K.
Tech. Director,
Foreign Tech. Div.
Air Force Syst. Cmd.

GARDINER
L. Keith
Chief, Internatl
Issues Division
Offc of Pol. Analysis
Central Intell. Agency

HACKER
Benjamin T.
RADM, U.S. Navy
Commander, U.S.
Military Enlistment
Processing Command

HATCH
Henry J.
BGEN, U.S. Army
Cdr/Div. Engineer
Army Engineering
Div., Pacific Ocean

HUCK
Daniel F.
Principal Analyst
Natl Security &
Internat'l Affairs
Cong. Budget Office

JEFFCOAT
Clyde E., Jr.
Deputy Commander,
Finance & Accnting
Dept. of Army

JENES
Theodore G., Jr.
MGEN, U.S. Army
Dep. Commander
U.S. Combined Arms
Combat Div. Activity

JOHNSTON
Oswald L., Jr.
Diplomatic
Correspondent,
L.A. Times

KINKEAD
William W.
Dir. Transportation &
Maritime Policy, Offc
of Asst. SECNAV for
Shipbldg & Logistics

LANGSTON
Joann H.
Director, Study
Prog. Mngment Offc
Dept. of Army

LANOUE
Robert J.
Acting Dir., NATO
Prog. & Foreign
Construction
Offc of Secy of Def.

LARSON
Charles R.
RADM, U.S. Navy
Director, Long Range
Planning Group
Offc of Chief of
Naval Operations

LYNN
Verne L.
Director, Defensive
Systems
Offc of Secy of Def.

McCARTHY
James P.
BGEN, U.S. Air Force
Special Assistant
for MX, DCS
R, D & Acquisition

McCAUSLAND
Charles
BGEN, U.S. Air Force
Dep. Chief of Staff
Plans & Programs
AF Logistics Cmd.

McFARLAND
Edward L.
BGEN, U.S. Air Force
Reserve
Mobilization Asst. to
Asst DCS, Plans & Ops

RIDGWAY
Rozanne L.
Special Asst. to
Secy of State for
Bilateral Negotiations
Dept. of State

ROBINSON
Clarence A.
Sr Military Editor
Aviation Week &
Space Technology

ROCKWELL
James M.
MGEN, U.S. Army
Dir. for Command
Control & Communic.
CINCPAC

ROSENBLUM
Howard E.
Deputy Director for
Communications
Security
Natl Security Agncy

RUBINO
D. Jerry
Staff Director
Security Programs
Dept. of Justice

SARBER
William R. Jr.
BGEN, U.S. Army
Asst. DCS Logistics
for Security Assist.
HQ, U.S. Army

SCHWANKL
Gerald C.
BGEN (sel.),
U.S. Air Force
Deputy, Armament
Div., AF Syst. Cmd.

SEIGLE
John W.
MGEN, U.S. Army
Director of Strategy,
Plans & Policy; Offc
DCS for Ops & Plans

TUCKER
Kevin M.
Special Agent
in Charge
Phil. Field Office
U.S. Secret Service

USHER
William R.
MGEN, U.S. Air Force
Personnel Director,
AF Personnel Plans

VANDER SCHAAF
Derek J.
Sr. Staff Advisor
House Appropriations
Committee

WHITE
Justus P.
Prof. Staff
House Armed
Services Committee

WILLARD
Richard K.
Counsel to Atty. Gen.
for Intell. Policy
Dept. of Justice

WYMAN
Parker D.
Internatl Affairs
Adviser, Industrial
College of the
Armed Forces

ZEIMER
Dale E.
Chief, New York
Field Office
Central Intell. Agency

MILLER
Arthur R.
Harvard Law School

MOORE
Jonathan
Kennedy School of
Government

NACHT
Michael L.
Kennedy School of
Government

NYE
Joseph S., Jr.
Kennedy School of
Government

PORTER
Roger B.
(on leave) Kennedy
School of Govt

PRICE
Don K.
Kennedy School of
Government

RAIFFA
Howard
Harvard Business
School

UYTERHOEVEN
Hugo E. R.
Harvard Business
School

Ratio scale ratings relating to
the perspectives and roles of other
participants in the national security process

(PLEASE READ CAREFULLY!)

Please evaluate how competent you feel in dealing with/operating in each of the listed areas. If you feel completely comfortable about operating in an area with which you are very familiar, a rating of 1,000 would be appropriate. If you only feel "half comfortable", then the rating becomes 500. If you feel one-tenth as comfortable (as compared to a very well-known area), then the rating should be 100. Ratings of zero (0) and negative ratings should not be used. Please assign a separate score to each Roman numeraled section (not the individual questions) for your pre-program and immediate post-program evaluations. This questionnaire will be returned to you in six months for similar evaluation at that time.

The questions listed within each section are intended to be representative of the types of things you should consider in deciding how comfortable you are. They should not be interpreted as representing an exhaustive treatment of any given area.

Before Program | Following Program | 6 mos later

RATIO SCALE RATINGS

I. The Presidency

- To what extent do Presidents embody national purpose and policy? To what extent are they leaders of parties or factions? To what extent do they act just for themselves? How can you tell? Can they?
- How do Presidents make decisions? Do they actually make them? Where do they get information: Given time constraints, how do they choose among competing presentations? How do they allocate scarce time? How do they determine priorities; or are their priorities determined by other people?
- How do Presidents differ? How important is the President's own background? How important are campaigns -- promises, etc.? How are Presidents different during and after transitions? . . . in second terms? What is the range of potential differences in operating style?
- Who else, if anyone, thinks Presidentially? How can or do Presidents deal with the fact that almost everyone else has a more compartmented outlook?
- Is the President just the big chief? Are his problems the problems of any executive or leader, or are they different in kind from problems of a cabinet officer, a corporation head, etc.?
- To what extent does the president's organization of his White House staff affect policy formulation and implementation?

II. The Congress

- What are the nuances of the legislative process that might make a difference in the way one frames one's programs prior to submission for Congressional approval?
- What are the priorities that are likely to be accorded competing demands for resources by the various national security-related Congressional committees?
- On what issues is Congressional decision-making likely to be driven by the national interest, as opposed to local political considerations?
- How should one deal with the situation where it looks likely that the substance of the issue may be subsumed in jurisdictional infighting, i.e., protection of turf vis-à-vis the Executive Branch, between committees, etc."
- To what extent should one provide one's personal opinion vis-à-vis total support for the Administration position?

III. Political Appointees

- How should one deal with the fact that the terms and conditions of employment of political appointees are considerably different from those for other categories of participants in the national security process?
- What motivates them?
- How does their different risk orientation affect their behavior and perception of things?
- What is the best way to cope with the managerial shortcomings of the various categories of political appointees? i.e., educators often are not skilled in the management of large organizations; lawyers are often not skilled in either quantitative analysis or the administration of large organizations; and businessmen often have trouble operating in a "no bottom line" environment.
- In what sorts of circumstances are political considerations likely to prevail over those of substance?

IV. Careerists

- To what extent does the career civilian's time horizon influence his or her approach to substantive decision-making?
- How should the career military man reconcile the often competing demands of loyalty to Service vis-a-vis the requirement to take direction from civilian officials?
- What are the differences in planning priorities between military and civilian planners?
- What are the effects of frequent turnover of military personnel on program decision-making and accountability?
- To what extent does a "civil service mentality" impede creativity and initiative in program formulation and implementation?

Before Program | Following Program | 6 mos later

___ ___ ___ V. Media

- How does one reconcile the inherent conflict between the media's interpretation of obligations and rights under the First Amendment and the need for secrecy that often underlies "national security considerations"?

- What motivates the press in its pursuit of news?

- How can one use the media to help achieve one's own objectives?

- How does one avoid press chicanery, such as the slanting of stories to achieve sensationalism and quoting out of context?

- What is the language with which one should be familiar when dealing with the media ("off the record", "deep background". "non-attribution", etc.)?

- Is citing a "personal opinion" ever appropriate for a top level official when dealing with the media?

___ ___ ___ VI. Private Sector

- What are the principal differences between public and private sector operating environments?

- What capabilities does an Administration have for orchestrating the disparate actions of multiple private entities in the national interest (such as the sale of wheat by U.S. farmers to other countries)?

- To what extent can the U.S. influence major international concerns (such as the major oil companies?

- How are trade-offs calculated when different values are involved (such as, when the price of more jobs includes a dirtier environment)?

Before Program | Following Program | 6 mos lat

___ ___ ___ VII. International

- In negotiating with a foreign government or in mediating between foreign governments, what approach can be taken to maximize the likelihood of working out terms that are mutually advantageous (as opposed to terms that represent victory for one party and defeat for the other)?

- When negotiating an agreement with representatives from another country, to what extent is it necessary/ desirable to understand not only their nation's real interests but the historical perspective in which they see the issues? Can it be done? Are there situations in which it would be better not to know at all rather than to know only a little?

- What are some of the key determinants of differences in national perspectives?

- What are some of the principal differences between American and Soviet approaches to strategic arms control?

_____

Signature

	PRES			CONGRESS			POLITICIANS			CAREERISTS			PRIV. SECTOR			INTERNATIONAL			*War	**Mil
Participants	Pre	Post	6 mo	Pre	Post	6 mo	Pre	Post	6 mo	Pre	Post	6 mo	Pre	Post	6 mo	Pre	Post	6 mo		
#1	800.	850.	750.	800.	800.	800.	800.	800.	800.	800.	800.	900.	500.	800.	900.	800.	800.	800.	700.	900.	900.	0.	0.
2	200.	700.	500.	500.	800.	800.	500.	600.	600.	800.	800.	900.	500.	700.	700.	800.	800.	800.	500.	800.	800.	0.	0.
3	600.	800.	900.	300.	700.	800.	700.	700.	800.	800.	800.	800.	750.	900.	900.	400.	700.	700.	400.	900.	950.	0.	0.
4	500.	700.	600.	500.	700.	800.	300.	500.	600.	300.	500.	500.	300.	700.	800.	100.	300.	300.	300.	700.	600.	0.	0.
5	400.	600.	600.	900.	900.	900.	400.	600.	600.	900.	900.	900.	500.	700.	700.	700.	800.	800.	900.	900.	900.	0.	0.
6	300.	600.	800.	400.	600.	900.	400.	750.	1000.	800.	900.	1000.	100.	500.	750.	500.	800.	900.	300.	600.	900.	0	1.
7	350.	600.	600.	200.	650.	800.	100.	600.	800.	800.	850.	900.	800.	850.	850.	100.	500.	700.	500.	850.	900.	1.	1.
8	300.	600.	700.	500.	700.	800.	700.	700.	700.	700.	700.	700.	100.	300.	300.	100.	500.	600.	100.	500.	500.	0.	0.
9	100.	500.	500.	500.	500.	500.	1000.	1000.	1000.	1000.	1000.	1000.	100.	100.	100.	100.	100.	100.	500.	500.	1000.	0.	0.
10	400.	600.	700.	200.	400.	500.	400.	650.	500.	600.	700.	600.	400.	500.	600.	500.	600.	650.	200.	300.	400.	0.	0.
11	200.	500.	700.	200.	400.	600.	100.	300.	500.	700.	800.	800.	300.	500.	600.	300.	600.	700.	400.	700.	700.	1.	1.
12	100.	400.	600.	200.	600.	800.	800.	800.	600.	900.	900.	900.	750.	850.	900.	700.	850.	900.	400.	600.	600.	0.	0.
13	400.	600.	700.	800.	800.	900.	700.	800.	800.	600.	800.	800.	600.	800.	800.	500.	700.	800.	500.	700.	800.	1.	0.
14	400.	600.	600.	500.	700.	850.	900.	900.	900.	800.	600.	800.	650.	650.	750.	700.	800.	850.	600.	800.	900.	1.	1.
15	200.	500.	700.	200.	400.	700.	100.	400.	550.	200.	500.	775.	200.	600.	700.	200.	500.	600.	300.	700.	750.	1.	1.
16	200.	600.	700.	100.	600.	700.	100.	500.	600.	400.	600.	800.	100.	200.	300.	100.	200.	400.	200.	400.	600.	0.	1.
17	500.	500.	700.	200.	300.	300.	300.	300.	400.	300.	100.	300.	500.	500.	500.	100.	100.	100.	500.	600.	600.	0.	0.
18	900.	950.	950.	600.	850.	850.	800.	850.	850.	700.	750.	750.	800.	900.	900.	600.	800.	850.	850.	900.	950.	0.	1.
19	500.	700.	900.	500.	700.	750.	700.	800.	600.	800.	900.	950.	900.	600.	800.	600.	700.	850.	800.	900.	950.	0.	0.
20	500.	700.	700.	500.	700.	750.	500.	800.	800.	500.	700.	750.	500.	800.	750.	800.	800.	800.	500.	650.	550.	1.	1.
21	600.	800.	900.	500.	800.	800.	400.	500.	500.	900.	900.	1000.	400.	800.	800.	500.	800.	800.	400.	700.	700.	1.	0.
22	600.	800.	800.	700.	800.	800.	700.	750.	700.	800.	850.	850.	200.	500.	300.	900.	900.	900.	500.	650.	550.	0.	0.
23	300.	600.	700.	500.	700.	700.	1000.	1000.	1000.	1000.	1000.	1000.	100.	500.	500.	400.	800.	800.	300.	600.	700.	0.	0.
24	800.	900.	900.	900.	950.	900.	600.	600.	700.	900.	900.	900.	700.	800.	800.	800.	900.	900.	300.	600.	600.	1.	1.
25	600.	800.	800.	200.	400.	400.	300.	500.	500.	600.	800.	900.	100.	500.	600.	300.	400.	500.	100.	400.	300.	0.	0.
26	500.	700.	800.	600.	650.	800.	500.	500.	600.	700.	700.	700.	300.	500.	600.	400.	700.	700.	700.	750.	800.	0.	1.
27	700.	800.	800.	400.	600.	600.	400.	500.	500.	500.	600.	600.	500.	600.	600.	600.	700.	700.	400.	700.	700.	0.	0.
28	700.	750.	800.	700.	700.	750.	700.	700.	750.	700.	700.	750.	700.	800.	800.	700.	700.	700.	700.	800.	800.	0.	0.
29	600.	800.	900.	800.	900.	900.	900.	900.	900.	900.	900.	1000.	400.	600.	700.	600.	800.	800.	500.	800.	800.	0.	0.
30	600.	800.	900.	800.	900.	900.	500.	750.	750.	800.	800.	800.	500.	800.	750.	700.	850.	850.	700.	850.	850.	1.	1.
31	600.	800.	800.	400.	600.	600.	500.	800.	800.	700.	800.	800.	400.	700.	700.	800.	700.	800.	600.	800.	900.	0.	1.
32	200.	600.	750.	300.	600.	650.	600.	700.	850.	700.	800.	750.	500.	600.	600.	100.	500.	650.	300.	600.	800.	0.	0.
33	600.	700.	750.	500.	600.	600.	600.	700.	700.	800.	850.	850.	300.	400.	500.	600.	650.	600.	400.	600.	600.	0.	0.
34	500.	750.	800.	500.	700.	800.	600.	800.	800.	450.	650.	700.	400.	600.	650.	500.	700.	700.	600.	800.	850.	0.	0.
35	500.	700.	700.	700.	800.	800.	700.	800.	700.	700.	700.	700.	700.	700.	800.	800.	400.	600.	500.	300.	600.	600.	0.	0.
36	500.	1000.	1000.	1000.	1000.	1000.	500.	1000.	1000.	100.	500.	700.	100.	500.	500.	100.	500.	500.	100.	500.	700.	0.	1.
37	500.	600.	650.	600.	700.	800.	400.	450.	600.	900.	950.	950.	400.	600.	650.	500.	600.	700.	900.	950.	950.	0.	0.
38	250.	750.	750.	500.	750.	900.	750.	900.	900.	1000.	1000.	1000.	250.	700.	800.	700.	900.	900.	250.	500.	500.	0.	0.
39	300.	600.	700.	700.	800.	900.	600.	800.	900.	500.	600.	800.	700.	700.	800.	600.	800.	800.	500.	800.	800.	0.	0.
40	200.	500.	500.	300.	600.	600.	400.	600.	700.	400.	300.	800.	500.	700.	700.	500.	700.	700.	200.	600.	700.	1.	1.
41	480.	660.	800.	400.	640.	750.	500.	820.	900.	560.	640.	800.	380.	700.	700.	350.	625.	725.	275.	600.	700.	1.	1.

*War: 0 = non-War College; 1 = War College

**Mil: 0 = Civilian; 1 = Military

Attachment F

INTRODUCTION TO THE T-TEST OF SIGNIFICANCE

In many investigations the researcher is primarily interested in discovering and evaluating differences between effects, rather than the effects themselves. For example, one may be interested in the difference in income for people at various levels of education. The most common of this type of analysis is the comparison of two groups of subjects, with the group means as the basis for comparison. An example of this would be to determine the difference in income between college graduates and nongraduates. This example is an instance where the two groups preexist the analysis. In some cases, a researcher may randomly assign subjects to two groups and apply a treatment to one group. Treatment effects are measured by comparing the two groups. For example, the effect of a brand of toothpaste on the prevention of cavities might be tested this way.

In the comparison of group means, the term *treatment* is used to refer to the basis on which the two groups are differentiated. In the first example, college education is the treatment; in the second example, it is very natural to call the toothpaste the treatment.

Statistical Inference from a Sample of Cases

Since it is most often impossible, or at least impractical, to compute a group mean based on all members of the group, the researcher must use a sample. The true but unknown mean for a group is called the *population mean;* it is estimated by the *sample mean*. The comparison of two group means is thus a problem of comparison of two sample means, and from that, inferring the difference between the means of the parent populations.

The basic problem is to determine whether or not a difference between two samples implies a true difference in the parent populations.

Since it is highly probable that two samples from the same population would be different due to the natural variability in the population, it is clear that a difference in sample means does not necessarily imply that the populations from which they were drawn actually differ on the characteristic being studied.

The goal of the statistical analysis is to establish whether or not a difference between two samples is significant. "Significant" here does not mean "important" or "consequence"; it is used here to mean "indicative of" or "signifying" a true difference between the two populations. The systematic approach used to test sample differences is as follows.

1 A null hypothesis and a corresponding alternative hypothesis are formulated. The null hypothesis (H_0) must be a precise statement for which the Student's t statistic (and probability) can be computed.

   Typically, H_0 is what the researcher is trying to disprove or reject so that the alternative hypothesis (H_1) can be accepted. Most often H_0 states that the population means are the same (H_0: $\mu_1=\mu_2$). Another possible statement for H_0 is that the population means differ by a specific amount, for example, H_0:$\mu_1-\mu_2=5.2$. H_1 is usually the set of all other possible outcomes (for example, H_1:$\mu_1\neq\mu_2$ or H_1:$\mu_1-\mu_2\neq5.2$).

2 A "significance level" for testing H_0 is chosen. Since sampling is being used, a decision to accept or reject H_0 cannot be made with absolute certainty; the decision must be based on probabilities. The significance level is the smallest probability that will be accepted as reasonable, i.e., due to chance or sample variability.

3 The samples are taken and the two sample means and variances are computed.

4 Assuming H_0 is true, the t statistic (see below) is computed. From the frequency distribution of the statistic is computed the probability of getting a more extreme value of the statistic. Intuitively, this is the probability of drawing two samples that differ more than the pair actually drawn.

5 If the probability computed in step 4 is smaller than the significance level chosen in step 2, H_0 is rejected. If the probability is larger, H_0 is not rejected. However, this does not necessarily imply that H_0 is true, only that the true situation is not significantly different from that assumed by the null hypothesis.

---

[1] The T-TEST procedure was developed, programmed, and documented for the CDC version of SPSS by James Tucci of Northwestern University

Typical values for the significance level chosen in step 2 are .05 or .01. The specific value of the significance level chosen is based on the seriousness of the type I error (rejecting H_0 when it is true) as opposed to type II error (accepting H_0 when it is false). The significance level is exactly the probability of rejecting H_0 when it is true. Thus, if type I error is very serious, the significance level would be set correspondingly low (.001 is sometimes used). On the other hand, if type II error has the worse consequence, the significance level could be raised, e.g., .10.

The *t* Distribution

Student's t is the statistic used in calculating the probability associated with H_0. The t is a statistic generally applicable to a normally distributed random variable where the mean is known (or as we shall see, assumed to be known) and the population variance is estimated from a sample. Assume that the normally distributed random variable X has mean μ and unknown population variance σ^2, which is estimated by a sample variance s^2. Then, $t = (X-\mu)s$. Note that this formula is almost identical to that for the standard normal deviate $z = (X-\mu)/\sigma$. Historically, the statistic z was always used for computing the probability of occurrence for normal variables, and tables of such probabilities were constructed and used for that purpose. The statistician William Seeley Gosset (who used the pseudonym "Student") realized that if z were computed using s based on a small sample, the use of normal tables was not trustworthy and an alternative table was needed. The t distribution is a direct result of his work, thus the name *Student's t*. The t distribution depends on the degrees of freedom used in computing s. This is the denominator in the s^2 calculation below and, in the case of one sample of size n, is $n - 1$ where

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

t is tabulated for various degrees of freedom, usually from 1 to 30. For degrees of freedom larger than 30, s^2 (the sample variance) is a sufficiently reliable estimate of σ^2 (the population variance) so that the distribution of t is almost identical to that of z. This is indicated in t tables by the entry ∞ (infinity) for degrees of freedom. The probability given for t is usually *two-tailed*, that is, the probability for a value of $|t|$ (the absolute value of t) or larger.

The t is a statistic which may be computed for a normally distributed variable; to compute a t value for a pair of sample means, the following points must be considered:

1 The sample mean is a normally distributed variable. That is, given a normal population with mean μ and variance of σ^2, if samples of size n are drawn, the sample means are normally distributed with mean μ and variance (σ^2/n).
2 The difference of two normally distributed random variables is a normal random variable. That is, given two random variables X_1 and X_2 with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively, the random variable $D = X_1 - X_2$ generated by pairs independently selected from the two populations is normally distributed with mean $\mu_1 - \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

From the above, the difference of sample means is a normally distributed variable with mean $(\mu_1 - \mu_2)$ and variance $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$.

Comparison of Means—Independent Samples, Populations with Common Variance

Given two populations with means μ_1 and μ_2, respectively, and common variance σ^2, all unknown, the problem is to use the t-test of significance to determine if $\mu_1 = \mu_2$. Following the steps outlined above:

1 The null hypothesis H_0:$\mu_1 = \mu_2$ and alternative hypothesis H_0:$\mu_1 \neq \mu_2$ are formulated.
2 The significance level α is chosen.
3 The two populations are sampled; means $\bar{x}_1$ and $\bar{x}_2$, variances s_1^2 and s_2^2 are computed, based on samples of sizes n_1 and n_2, respectively. From the pair of sample variances, which are both estimates of σ^2, the "pooled variance" is computed.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

[Note that there are $(n_1 + n_2 - 2)$ degrees of freedom in the computation.] s^2 is the weighted average of the sample variances and is the best estimator for σ^2. Writing $\bar{d} = \bar{x}_1 - \bar{x}_2$, since the population variance for the difference in sample means is $\sigma_{\bar{d}}^2 = \sigma_1^2/n_1 + \sigma_2^2/n_2$, the sample variance for the difference of sample means is

$$s_{\bar{d}}^2 = (s^2/n_1 + s^2/n_2)$$

t corresponding to the difference in sample means is computed:

$$t_{\bar{d}} = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_1 - \mu_2}{s_{\bar{d}}}$$

which, under the null hypothesis H_0: $(\mu_1 = \mu_2)$, reduces to

$$t_{\bar{d}} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{\bar{d}}} \quad \text{with } (n_1 + n_2 - 2) \text{ degrees of freedom.}$$

4 The probability associated with t is computed. The probability is for the occurrence of a value equal to or larger than t, sign ignored. This is the two tailed probability and it is appropriate to the set of hypotheses chosen since they do not assume that t will be either positive or negative. Tests of this type, i.e., where H_1 specifies inequality, are called *two-tailed tests*. One-tailed tests may also be used, as described in Sec. 17.2.1.6.
5 H_0 is rejected if the two-tailed probability for H_0 is less than α chosen in step 2.

Comparison of Means—Independent Samples, Populations with Unequal Variances

Given populations with unequal variances, t *cannot* be computed for the difference in sample means. Instead, an approximation to t may be computed.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_1^2/n_1 + s_2^2/n_2}$$

This statistic is not distributed as Student's t. However, the probability for t can be approximated by treating it as t, but with degrees of freedom

$$\text{df} = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

This number is usually not an integer, but a reasonably accurate probability is obtained by rounding to the nearest integer.

If it is not known whether the two populations have the same variance, an F test of sample variances may be performed: The null hypothesis $H_0{:}\sigma_1^2 = \sigma_2^2$ with alternative $H_1{:}\sigma_1^2 \neq \sigma_2^2$ and a significance level α' is chosen (α' does not necessarily have the same value as α used for the t-test). From the sample variances, F is computed.

$$F = \frac{\text{larger } s^2}{\text{smaller } s^2}$$

If the probability for F is greater than α', H_0 is accepted; t based on the pooled-variance estimate for σ_D^2 should be issued.

If the probability for F is less than or equal to α', H_0 is rejected; t based on the separate variance estimate for σ_D^2 should be used.

CLUSTER ANALYSIS

The specific algorithm used for the cluster analysis operates as follows:

1) Data Set: the data used consists of the 7 "during-program" gains for each of 30 participants.

2) First compute the means, $\bar{x}_i$, and standard deviations, s_i, for each area, where i indexes the area.

3) Standardize each participant's score. That is, if x_i equals a given participant's score for area i, then his standardized score (sometimes referred to as a z-score) for that area equals $\frac{x_i - \bar{x}_i}{s_i}$. This is done for each of the seven areas for all of the cases.

4) Compute the distance between each case. The distance is defined as follows: Let y_i, i=1, 7 be the standardized scores for case y and z_i, i=1, 7 for case z. Then the distance between y and z is equal to

$$\sqrt{\sum_{i=1}^{7}(y_i - z_i)^2} .$$

This distance, d, is computed for all pairwise combinations of cases.

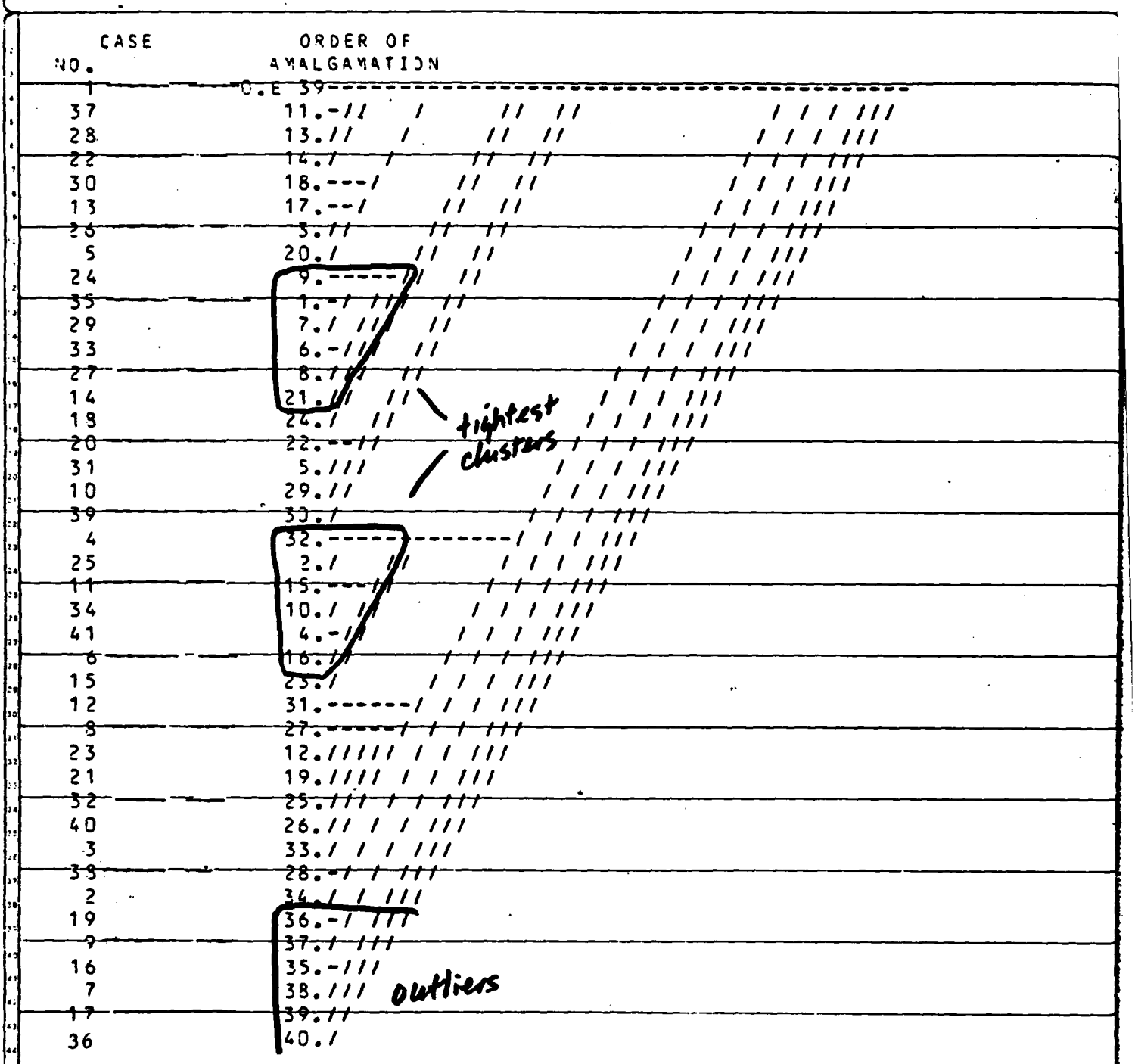
5) Begin forming clusters by joining the two cases with the smallest distance between them. Note that this distance is a

measure of how similar cases are; the more alike the cases (y_i-z_i will be closer to 0), the smaller is the distance between them. Where two cases are joined, the average of each y_i and z_i [=1/2(y_i+z_i)] becomes the standardized scores for the cluster.

6) Update the distance lists by replacing the distances from each case to cases y and z with the distance to the yz cluster.

7) Continue joining the cases into clusters; at each step join either the two closest cases, the two closest clusters or the closest case-to-cluster pair, whichever is the smallest distance overall. After each joining, update the distance lists so that the new distances reflect the weighted average of the elements combined in the new cluster. (Note: keep track of the distance between the cases or clusters that are joined.)

8) Stop clustering when all cases are joined into a single large cluster. (This is a simple, straightforward cluster technique used in the BMDP2M program of the Biomedical Computer Program P-Series (1977) developed at the Health Sciences Computer Facility, UCLA. For more complex, specialized, albeit sophisticated and powerful techniques, see Clustering Algorithms, Hartigan, J.A., John Wiley & Sons, New York, 1975).


CASE NO.
ORDER OF AMALGAMATION
O.E 39
11.-// / // // / / / ///
13.// / // // / / / ///
14./ / // // / / / ///
18.---/ // // / / / ///
17.--/ // // / / / ///
3.// // // / / / ///
20./ // // / / / ///
9.-----// // / / / ///
1.-/ /// // / / / ///
7./ /// // / / / ///
6.-/// // / / / ///
8./ // // / / / ///
21.// // / / / ///
24./ // / / / ///
22.--// / / / ///
5./// / / / ///
29.// / / / ///
30./ / / / ///
32.-----/ / / ///
2./ /// / / / ///
15.---// / / / ///
10./ // / / / ///
4.-// / / / ///
16./ / / / ///
25./ / / / ///
31.------/ / / ///
27.-----/ / / ///
12.///// / / ///
19.//// / / ///
25./// / / ///
26.// / / ///
33./ / / ///
28.-/ / ///
34./ / ///
36.-/ ///
37./ ///
35.-///
38./// outliers
39.//
40./
tightest clusters

CORRELATION ANALYSIS

1) $\bar{x}$ and $\bar{y}$ are the average "during-program" gains in areas x and y.

2) x_i and y_i are the gains of case i in areas x and y.

3) Var x and Var y are the variances of "during-program" gains in areas x and y.

4) Using the above definition, $$\rho = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})\,(y_i-\bar{y})}{\sqrt{Var(x)Var(y)}}$$

where N is the number of participants and ρ, the "correlation coefficient", varies between -1 and 1. When x and y are perfectly correlated, that is, when x is a positive linear function of y, or they both are linear functions (with the same sign) of some other variable z, then $\rho=1$. If x is a negative linear function of y, then $\rho=-1$. If x and y are uncorrelated, $\rho=0$. In general terms, if $\rho>0$, x increases as y increases, and if $\rho<0$, x decreases as y increases. The closer to 1 (or -1) ρ is, the more closely x tracks changes in y.

The application to the problem at hand is as follows: If an attendee gains a great deal during the program in area A, can he or she be expected to gain significantly in area B?